

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2004, 14:19:41 ; Search time 64 Seconds
(without alignments)
3964.495 Million cell updates/sec

Title: US-10-624-932-2

Perfect score: 4791

Sequence: 1 MAVRPGLWPALLGIVLAAWL.....AVAGLGQPDAGLFTVSEAE 898

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:
1: geneseqp1980s:
2: geneseqp1990s:
3: geneseqp2000s:
4: geneseqp2001s:
5: geneseqp2002s:
6: geneseqp2003as:
7: geneseqp2003bs:
8: geneseqp2004s:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			Description
		Match	Length	DB ID	
1	4791	100.0	898	5 AAU85403	Aau85403 Human pro
2	4781	99.8	898	5 AAU97899	Aau97899 Human net
3	4698.5	98.1	899	5 AAU79939	Aau79939 Human UNC
4	4638	96.8	898	2 AAW78898	Aaw78898 Rat UNC-5
5	4638	96.8	898	5 AAU10543	Aau10543 Rat netri
6	4638	96.8	898	5 AAU97900	Aau97900 Rat netri
7	4526.5	94.5	943	4 AAM79128	Aam79128 Human pro
8	4413	92.1	842	5 AAU74818	Aau74818 Human REP
9	2815	58.8	556	2 AAW78899	Aaw78899 Human UNC

10	2755	57.5	931	4	AAB50691	Aab50691 Human UNC
11	2755	57.5	931	7	ADE63098	Ade63098 Human Pro
12	2755	57.5	982	4	ABG11551	Abg11551 Novel hum
13	2578.5	53.8	945	7	ADE63096	Ade63096 Rat Prote
14	2571.5	53.7	943	2	AAW78900	Aaw78900 Rat UNC-5
15	2563.5	53.5	933	5	AAO18734	Aao18734 Human NOV
16	2563.5	53.5	933	5	AAO18735	Aao18735 Human NOV
17	2558.5	53.4	945	4	AAU12244	Aau12244 Human PRO
18	2558.5	53.4	945	6	ABO17688	Abo17688 Novel hum
19	2558.5	53.4	945	6	ABU80942	Abu80942 Human PRO
20	2558.5	53.4	945	6	ABU66642	Abu66642 Human PRO
21	2558.5	53.4	945	6	ABU59723	Abu59723 Novel sec
22	2558.5	53.4	945	6	ABO24913	Abo24913 Human sec
23	2558.5	53.4	945	6	ABU66918	Abu66918 Human sec
24	2558.5	53.4	945	6	ADA45665	Ada45665 Novel hum
25	2558.5	53.4	945	6	ADA76096	Ada76096 Human PRO
26	2558.5	53.4	945	6	ADA18746	Ada18746 Human PRO
27	2558.5	53.4	945	6	ADA61369	Ada61369 Homo sapi
28	2558.5	53.4	945	6	ADB19154	Adb19154 Novel hum
29	2558.5	53.4	945	6	ADB27695	Adb27695 Human PRO
30	2558.5	53.4	945	6	ADA86174	Ada86174 Novel hum
31	2558.5	53.4	945	6	ADB15738	Adb15738 Human PRO
32	2558.5	53.4	945	6	ADA47524	Ada47524 Human PRO
33	2558.5	53.4	945	6	ADA67319	Ada67319 Human PRO
34	2558.5	53.4	945	6	ADB30326	Adb30326 Human PRO
35	2558.5	53.4	945	6	ADA85622	Ada85622 Novel hum
36	2558.5	53.4	945	6	ADA96834	Ada96834 Human PRO
37	2558.5	53.4	945	6	ADA79138	Ada79138 Human PRO
38	2558.5	53.4	945	6	ADA87277	Ada87277 Novel hum
39	2558.5	53.4	945	6	ADB16479	Adb16479 Human PRO
40	2558.5	53.4	945	6	ADA91571	Ada91571 Novel hum
41	2558.5	53.4	945	6	ADB14634	Adb14634 Human PRO
42	2558.5	53.4	945	6	ADB18595	Adb18595 Novel hum
43	2558.5	53.4	945	6	ADA93810	Ada93810 Human PRO
44	2558.5	53.4	945	6	ADB19706	Adb19706 Novel hum
45	2558.5	53.4	945	6	ADB13018	Adb13018 Human PRO

ALIGNMENTS

RESULT 1
 AAU85403
 ID AAU85403 standard; protein; 898 AA.
 XX
 AC AAU85403;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE Human protein NOV1.
 XX
 KW Human; NOVX; cardiomyopathy; atherosclerosis; diabetes;
 KW cell signal processing disorder; metabolic disorder; obesity; infection;
 KW anorexia; cancer-associated cachexia; cancer; neurodegenerative disorder;
 KW Alzheimer's disease; Parkinson's disease; immune disorder;
 KW haematopoietic disorders; dyslipidaemia; pain; asthma; hypertension;
 KW osteoporosis; Crohn's disease; multiple sclerosis; angina pectoris;

KW myocardial infarction; ulcer; allergy; benign prostatic hypertrophy;
KW psychosis; neurological disorder; anxiety; schizophrenia;
KW manic depression; dementia; dyskinesia; Huntington's disease;
KW Gilles de la Tourette's syndrome; gene therapy.

XX

OS Homo sapiens.

XX

PN WO200210216-A2.

XX

PD 07-FEB-2002.

XX

PF 30-JUL-2001; 2001WO-US024225.

XX

PR 28-JUL-2000; 2000US-0221409P.

PR 04-AUG-2000; 2000US-0222840P.

PR 04-AUG-2000; 2000US-0223752P.

PR 04-AUG-2000; 2000US-0223762P.

PR 04-AUG-2000; 2000US-0223769P.

PR 04-AUG-2000; 2000US-0223770P.

PR 14-AUG-2000; 2000US-0225146P.

PR 15-AUG-2000; 2000US-0225392P.

PR 15-AUG-2000; 2000US-0225470P.

PR 16-AUG-2000; 2000US-0225697P.

PR 01-FEB-2001; 2001US-0263662P.

PR 05-APR-2001; 2001US-0281645P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Padigaru M, Mezes P, Mishra V, Burgess C, Casman S, Grosse WM;

PI Alsobrook JP, Lepley DM, Gerlach VL, Macdougall JR, Smithson G;

XX

DR WPI; 2002-180074/23.

DR N-PSDB; ABK37922.

XX

PT New isolated cytoplasmic, nuclear, membrane bound, or secreted polypeptide, useful for treating cardiomyopathy, atherosclerosis, infections, cancer, neurodegenerative, metabolic, hematopoietic and immune disorders.

XX

PS Claim 1; Page 11; 213pp; English.

XX

CC The invention relates to an isolated cytoplasmic, nuclear, membrane bound, or secreted polypeptide (NOVX, x= 1-14) their variants or mature form. Also included are the nucleic acids encoding the NOVX proteins, a vector comprising the nucleic acid, a cell comprising the vector, an anti-NOVX antibody and modulators of NOVX. NOVX, the nucleic acid and the antibody are useful for treating or preventing a NOVX-associated disorder, where the disorder is selected from cardiomyopathy, atherosclerosis, diabetes, a disorder related to cell signal processing and metabolic pathway modulation, metabolic disorders, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's disease, Parkinson's disease, immune disorders, haematopoietic disorders, and the various dyslipidaemias, metabolic disturbances associated with obesity, the metabolic syndrome X and wasting disorders associated with chronic diseases, bacterial, fungal, protozoal and viral infections, pain, bulimia, asthma, hypertension, urinary retention, osteoporosis, Crohn's

CC disease, multiple sclerosis, Albright Hereditary Osteodystrophy, angina
CC pectoris, myocardial infarction, ulcer, allergy, benign prostatic
CC hypertrophy, and psychotic and neurological disorders, including anxiety,
CC schizophrenia, manic depression, delirium, dementia, and dyskinesias,
CC such as Huntington's disease and Gilles de la Tourette's syndrome. The
CC nucleic acid is useful in gene therapy. The present sequence represents a
CC NOVX protein

XX

SQ Sequence 898 AA;

Query Match 100.0%; Score 4791; DB 5; Length 898;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 898; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

QY 61 VLLVCKAVPATQIFFKCNGEWRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVFGL 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 VLLVCKAVPATQIFFKCNGEWRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVFGL 120

QY 121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

QY 181 VEWLRNEDLVPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 VEWLRNEDLVPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240

QY 241 VNGGWSTWTEWSVCASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGS 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 VNGGWSTWTEWSVCASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGS 300

QY 301 WSPWSKWSACGLDCTHWRSRECSDPAPRNGGECCQGTLDTRNCTS DLCVHSASGPEDVA 360
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 WSPWSKWSACGLDCTHWRSRECSDPAPRNGGECCQGTLDTRNCTS DLCVHSASGPEDVA 360

QY 361 LYVGLIAVAVCLVLLVLILVYCRKKEGLSDVADSSILTSGFQPVS IKPSKADNP HLL 420
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 LYVGLIAVAVCLVLLVLILVYCRKKEGLSDVADSSILTSGFQPVS IKPSKADNP HLL 420

QY 421 TIQPDLSTTTTYQGS LCP RQDG P S P K F Q L T N G H L L S P L G G R H T L H H S S P T S E A E E F V S 480
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 421 TIQPDLSTTTTYQGS LCP RQDG P S P K F Q L T N G H L L S P L G G R H T L H H S S P T S E A E E F V S 480

QY 481 RLSTQNYFRSLPRGTSNM TYGTFNFLGGRLMIPNTGISLLIPPA I PRGKIYEIYLTLHK 540
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 481 RLSTQNYFRSLPRGTSNM TYGTFNFLGGRLMIPNTGISLLIPPA I PRGKIYEIYLTLHK 540

QY 541 PEDVRLPLAGCQTLLS PIVSCGPPGVLLTRPVILAMDHC GEPSPDSW SLRLKKQS CEGSW 600
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 541 PEDVRLPLAGCQTLLS PIVSCGPPGVLLTRPVILAMDHC GEPSPDSW SLRLKKQS CEGSW 600

QY 601 EDVLHLGEEAPSHLYYCQLEASACYVFT EQLGRFALVGEALS VAAKRLKLLL FAPVACT 660
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 601 EDVLHLGEEAPSHLYYCQLEASACYVFT EQLGRFALVGEALS VAAKRLKLLL FAPVACT 660

QY 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVSSLW 720
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVSSLW 720

QY 721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDDLACKLWWQVEGDGQSFSINF 780
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDDLACKLWWQVEGDGQSFSINF 780

QY 781 NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDPPCRRGADWRTLAQKL 840
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 781 NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDPPCRRGADWRTLAQKL 840

QY 841 HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 898
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 841 HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 898

RESULT 2

AAU97899

ID AAU97899 standard; protein; 898 AA.

XX

AC AAU97899;

XX

DT 27-AUG-2002 (first entry)

XX

DE Human netrin binding membrane receptor UNC5H-1 protein.

XX

KW Netrin binding membrane receptor; receptor; UNC5H-1; human; nootropic;
KW neuroprotective; cytostatic; antiparkinsonian; cerebroprotective; cancer;
KW central nervous system; CNS; stroke; Parkinson's disease;
KW multiple sclerosis; Alzheimer's disease.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers
FT Domain 152. .223
FT /note= "Immunoglobulin domain "
FT Domain 247. .294
FT /note= "Thrombospondine type 1 domain "
FT Domain 302. .348
FT /note= "Thrombospondine type 1 domain"
FT Region 361. .382
FT /note= "Transmembrane region"
FT Domain 495. .598
FT /note= "ZU5 domain"
FT Domain 817. .897
FT /note= "Death domain"

XX

PN WO200233080-A2.

XX

PD 25-APR-2002.

XX

PF 15-OCT-2001; 2001WO-EP011891.

XX

PR 16-OCT-2000; 2000US-0240061P.

XX

PA (FARB) BAYER AG.

xx

PI Koehler RH;

xx

DR WPI; 2002-463314/49.

DR N-PSDB; ABK52891.

xx

PT Novel human netrin binding membrane receptor polypeptide and
PT polynucleotides for identifying modulating agents useful in treating
PT diseases e.g. Parkinson's disease, multiple sclerosis, stroke,
PT Alzheimer's disease.

xx

PS Claim 1; Fig 2; 94pp; English.

xx

This invention relates to the DNA and protein sequences of a novel purified human netrin binding membrane receptor, UNC5H-1. The DNA sequence of the invention is useful as a probe for detecting a nucleic acid encoding the UNC5H-1 protein in a biological sample. The sequences of the invention are useful to screen for agents which decrease the activity of the UNC5H-1 protein. The sequences are also useful for screening agents which regulate (modulate) the activity of the protein of the invention. A pharmaceutical composition containing the protein of the invention or a reagent that modulates the activity of the UNC5H-1 protein may be useful for treating a UNC5H-1 dysfunction related disease such as cancer or a central nervous system (CNS) disorders (e.g., Parkinson's disease, multiple sclerosis, stroke and Alzheimer's disease). Fusion proteins comprising the UNC5H-1 protein are useful for generating antibodies and for in various assay systems, and the protein can be used as a bait protein in a two-hybrid assay or three-hybrid assay. The method of the invention is useful for detecting a coding sequence for the UNC5H-1 protein. The present sequence represents the human netrin binding membrane receptor UNC5H-1 protein of the invention.

xx

SQ Sequence 898 AA;

Query Match 99.8%; Score 4781; DB 5; Length 898;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 896; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAVRPGLWPALLGIVLAALWRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKKNP 60

QY 61 VLLVCKAVPATQIFFKCNGEWRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVFGL 120

Qy 121 EYWCQCVAWSSSGTTSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAE 180

Qy 181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVIVY 240

Qy 241 VNGGWSTWTEWSVCSASCGRGWQKRSRSTNPAPLNGGAFCEGQNVQKTACATLCPVDGS 300

Db 241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGS 300
 Qy 301 WSPWSKWSACGLDCTHWRRECSDPAPRNGGEECQGTLDTRNCTS DL CVHSASGPEDVA 360
 |||||||
 Db 301 WSPWSKWSACGLDCTHWRRECSDPAPRNGGEECQGTLDTRNCTS DL CVHTASGPEDVA 360
 |||||||:
 Qy 361 LYVGLIAAVAVCLVLLLV LILVYCRKKEGLDSVADSSILTSGFQPVS IKPSKADNP HLL 420
 |||||||
 Db 361 LYVGLIAAVAVCLVLLLV LILVYCRKKEGLDSVADSSILTSGFQPVS IKPSKADNP HLL 420
 |||||||
 Qy 421 TIQPDLSTTTTYQGSLCPRQDGSPKFQLTNGHLLSPLGGGRH TLHHSSPTSEAE FVS 480
 |||||||
 Db 421 TIQPDLSTTTTYQGSLCPRQDGSPKFQLTNGHLLSPLGGGRH TLHHSSPTSEAE FVS 480
 |||||||
 Qy 481 RLSTQNYFRSLPRGTSNM TYGTFNFLGGRLMIPNTGISL LIPPDAI PRGKIYEIYLTLHK 540
 |||||||
 Db 481 RLSTQNYFRSLPRGTSNM TYGTFNFLGGRLMIPNTGISL LIPPDAI PRGKIYEIYLTLHK 540
 |||||||
 Qy 541 PEDVRLPLAGCQTLLS PIVSCGPPGVLLTRPVILAMDHC GEPS PDWSLRLKKQS CEGSW 600
 |||||||
 Db 541 PEDVRLPLAGCQTLLS PIVSCGPPGVLLTRPVILAMDHC GEPS PDWSLRLKKQS CEGSW 600
 |||||||
 Qy 601 EDVLHLGEAPSHLYYCQ LEASACYVFT EQLGRFALVGEALS VAAKRLKLLL FAPVACT 660
 |||||||
 Db 601 EDVLHLGEAPSHLYYCQ LEASACYVFT EQLGRFALVGEALS VAAKRLKLLL FAPVACT 660
 |||||||
 Qy 661 SLEYNIRVYCLHDTHDALKEV VQLEKQLGGQLI QEP RVLHF KDSYHNLRLSI HDVPSSLW 720
 |||||||
 Db 661 SLEYNIRVYCLHDTHDALKEV VQLEKQLGGQLI QEP RVLHF KDSYHNLRLSI HDVPSSLW 720
 |||||||
 Qy 721 KSKLLVSYQEIPFYHIWNGTQR YLHCTFTLERVSPSTS DLACKLWVWQVEGDGQSFSINF 780
 |||||||
 Db 721 KSKLLVSYQEIPFYHIWNGTQR YLHCTFTLERVSPSTS DLACKLWVWQVEGDGQSFSINF 780
 |||||||
 Qy 781 NITKDTRFAELLALESEAGVPA LVGPSAFKIPFLIRQKIISSLDP CRRGADWRTLAQKL 840
 |||||||
 Db 781 NITKDTRFAELLALESEAGVPA LVGPSAFKIPFLIRQKIISSLDP CRRGADWRTLAQKL 840
 |||||||
 Qy 841 HLD SHLSFFASKPSPTAMI LN LWEARHFPNGNLSQLAAA VAGLGQPDAGLFTVSEAEC 898
 |||||||
 Db 841 HLD SHLSFFASKPSPTAMI LN LWEARHFPNGNLSQLAAA VAGLGQPDAGLFTVSEAEC 898

RESULT 3

AAU79939

ID AAU79939 standard; protein; 899 AA.

XX

AC AAU79939;

XX

DT 15-JUL-2002 (first entry)

XX

DE Human UNC5-like protein NOV1.

XX

KW Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;
 KW cell signal processing; metabolic pathway modulation; cancerous tissue;
 KW antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;
 KW chromosome 13.

XX
OS Homo sapiens.
XX
PN WO200229038-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031377.
XX
PR 04-OCT-2000; 2000US-0237862P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Herrmann JL, Rastelli L, Shimkets RA;
XX
DR WPI; 2002-340104/37.
DR N-PSDB; ABK49422.
XX
PT Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for
PT treating cardiomyopathy, atherosclerosis, and cancer.
XX
PS Claim 1; Page 9; 180pp; English.
XX
CC The present invention relates to a new NOVX polypeptide having a 900
CC (NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 865 (NOV5), or 331 (NOV6)
CC residue amino acid sequence, as given in the specification. The novel
CC polypeptide, and its encoding polynucleotide, are used to treat
CC cardiomyopathy, atherosclerosis, cancer or a disease related to cell
CC signal processing and metabolic pathway modulation, in a human. Detecting
CC the polypeptide or polynucleotide is useful for identifying cancerous
CC tissue. The antibody can be used to treat diabetes or cancer. The host
CC cells can be used to produce non-human transgenic animals useful in drug
CC screening. The present amino acid sequence is that of the human UNC5-like
CC protein NOV1 of the invention. This sequence is encoded by the human UNC5
CC -like NOV1 gene located on chromosome 13
XX
SQ Sequence 899 AA;

Query Match 98.1%; Score 4698.5; DB 5; Length 899;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 888; Conservative 2; Mismatches 7; Indels 3; Gaps 3;

Qy 1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
Db 1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy 61 VLLVCKAVPATQIFFCNGEWRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVGLE 120
Db 61 VLLVCKAVPATQIFFCNGEWRQVDHVIERSTDGSSGEPTMEVRINVSRRQQVEKVGLE 120

Qy 121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
Db 121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy 181 VEWLRNEDLVDPSDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVIVY 240
Db 181 VEWLRNEDLVDPSDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVIVY 240

Qy	241	VNGGWSTWTEWSVCASCGRGWQKRSRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG	299
			: :
Db	241	VNGGWSTWTEWSVCASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVHDRTVSSLVSDG	300
Qy	300	SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEQCQGTDLDTRNCTS DLCVHSASGPEDV	359
Db	301	SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEQCQGTDLDTRNCTS DLCVHSASGPEDV	360
Qy	360	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSVADSSILTSGFQPVSIKPSKADNPHL	419
Db	361	ALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSVADSSILTSGFQPVSIKPSKADNPHL	420
Qy	420	LTIQPDLS-TTTTYQGSCLCPRQDGSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFFV	479
Db	421	LTIQPDLS-TTTTYQGSCLCPRQDGSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFFV	479
Qy	480	SRLSTQNYFRSLPRGTSNMTYGTNFLLGGLMI PNTGISLLIPPDAI PRGKIYEIYLTLH	539
Db	480	SRLSTQNYFRSLPRGTSNMTYGTNFLLGGLMI PNTGISLLIPPDAI PRGKIYEIYLTLH	539
Qy	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGS	599
Db	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGS	599
Qy	600	WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLL FAPVA	658
Db	600	WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLL FAPVA	659
Qy	659	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKD SYHNLRLSIHDVPSS	718
Db	660	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKD SYHNLRLSIHDVPSS	719
Qy	719	LWKS KL LVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS DLACKLWVWQVEGDGQSF SI	778
Db	720	LWKS KL LVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS DLACKLWVWQVEGDGQSF SI	779
Qy	779	NFNITKDTRFAELLA ESEAGVPA L VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ	838
Db	780	NFNITKDTRFAELLA ESEAGVPA L VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ	839
Qy	839	KLHLD SHLSFFASKPSPTAMI LN LWEARHFPNGNLSQLAAA VAGLGQPDAGLFTVSEAEC	898
Db	840	KLHLD SHLSFFASKPSPTAMI LN LWEARHFPNGNLSQLAAA VAGLGQPDAGLFTVSEAEC	899

RESULT 4

AAW78898

ID AAW78898 standard; protein; 898 AA.

XX

AC AAW78898;

XX

DT 25-MAR-2003 (revised)

DT 21-DEC-1998 (first entry)

XX

DE Rat UNC-5 homologue UNC5H-1.

XX

KW UNC-5; UNC5H-1; rat; netrin receptor; cell migration; axon guidance; diagnosis; therapy.
XX
OS Rattus sp.
XX
FH Key Location/Qualifiers
FT Peptide 580. .594
FT /note= "peptide used to raise rabbit polyclonal antisera"
XX
PN WO9837085-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-US003143.
XX
PR 19-FEB-1997; 97US-00808982.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Tessier-Lavigne M, Leonardo ED, Hinck L, Masu M, Keinomasu K;
XX
DR WPI; 1998-495364/42.
DR N-PSDB; AAV52940.
XX
PT Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and
PT the biopharmaceutical industry.
XX
PS Claim 1; Page 19-22; 32pp; English.
XX
CC UNC5H-1 and UNC5H-2 (see AAW78900) are rat homologues of *Caenorhabditis*
CC elegans UNC-5 protein. Their amino acid sequences were deduced from
CC isolated unc5h cDNA clones (see AAV52940 and AAV52942) isolated from an
CC E18 brain cDNA library. The predicted proteins show similarity with UNC-
CC 5, possess 2 predicted Ig-like domains and 2 predicted thrombospondin
CC type-1 repeats, a predicted membrane spanning region, and a large
CC intracellular domain. They are predicted to be involved in cell migration
CC and axon guidance, and are characterised as receptor proteins for
CC netrins. Human UNC5H-1 (see AAW78899) and UNC5H-2 (see AAW78901) proteins
CC are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly
CC from transfected host cells. The invention also provides unc-5
CC hybridisation probes and primers, vertebrate UNC-5-specific binding
CC agents such as specific antibodies, and methods of making and using the
CC subject compositions in diagnosis (e.g. genetic hybridisation screens for
CC vertebrate unc-5 transcripts), therapy (e.g. gene therapy to modulate
CC vertebrate unc-5 gene expression) and in the biopharmaceutical industry
CC (e.g. as immunogens, reagents for modulating cell guidance, reagents for
CC screening chemical libraries for lead pharmacological agents, etc.).
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 898 AA:

Query Match 96.8%; Score 4638; DB 2; Length 898;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 862; Conservative 17; Mismatches 19; Indels 0; Gaps 0;

Qy 1 MAVRPGLWPALLGIVLAAWI.RGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Db 1 MAVRPGILWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
Qy 61 VLLVCKAVPATQIFFKCNGEWRQVDHIERSTDGSSGLPTMEVRINVSRRQQVEKVFGLE 120
Db 61 VLLVCKAVPATQIFFKCNGEWRQVDHIERSTDSSGLPTMEVRINVSRRQQVEKVFGLE 120
Qy 121 EYWCQCVAWSSSGTTSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
Db 121 EYWCQCVAWSSSGTTSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
Qy 181 VEWLRNEDLVDPSDLPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVIVY 240
Db 181 VEWLRNEDLVDPSDLPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSTSAIVIVY 240
Qy 241 VNNGWSTWTEWSVCASCGRGWQKRSRSCTNPAPLNGGAFCEGQNQVKTACATLCPVDGS 300
Db 241 VNNGWSTWTEWSVCASCGRGWQKRSRSCTNPAPLNGGAFCEGQNQVKTACATLCPVDGS 300
Qy 301 WSPWSKWSACGLDCTHWRSRECSDPAPRNGGECCQGTDLDTRNCTS DLCVHSASGPEDVA 360
Db 301 WSSWSKWSACGLDCTHWRSRECSDPAPRNGGECCRGAIDLDRNCTS DLCCLHTASCPEDVA 360
Qy 361 LYVGLIAVAVCLLLLLVLILVYCRKKEGLDSVADSSILTSGFQPVSIKPSKADNPHELL 420
Db 361 LYIGLVAVAVCLFLLLALGLIYCRKKEGLDSVADSSILTSGFQPVSIKPSKADNPHELL 420
Qy 421 TIQPDLSTTTTYQGSCLCPRQDGSPKFQLTNGHLLSPLGGRHTLHHSSPTSEAEEFVS 480
Db 421 TIQPDLSTTTTYQGSCLCSRQDGSPKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEDFVS 480
Qy 481 RLSTQNYFRSLPRGTSNMTRYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHK 540
Db 481 RLSTQNYFRSLPRGTSNMAYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHK 540
Qy 541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSW 600
Db 541 PEDVRLPLAGCQTLLSPVVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSW 600
Qy 601 EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLFAPVACT 660
Db 601 EDVLHLGEEESPShLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLLFApVACT 660
Qy 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW 720
Db 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW 720
Qy 721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS DLACKLWWVQVEGDGQSFSINF 780
Db 721 KSKLLVSYQEIPFYHIWNGTQOQLHCTFTLERINASTS DLACKVWWVQVEGDGQSFSNF 780
Qy 781 NITKDTRFAELLALESEAGVPAVGPSAFKIPFLIRQKIISSLPPCRRGADWRTLAQKL 840
Db 781 NITKDTRFAELLALESEGGVPAVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQKL 840
Qy 841 HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAGLQPDAGLFTVSEAEC 898
Db 841 HLDSHLSFFASKPSPTAMILNLWEARHFPNGNGLGOLAAVAGLQPDAGLFTVSEAEC 898

RESULT 5
AAU10543
ID AAU10543 standard; protein; 898 AA.
XX
AC AAU10543;
XX
DT 14-FEB-2002 (first entry)
XX
DE Rat netrin receptor UNC5H1 (YSG7) polypeptide.
XX
KW YSG; YSG7; schizophrenia; chronic animal model; LCGU; netrin receptor;
KW local cerebral glucose utilisation; phosphodiesterase 1-alpha; UNC5H1;
KW calcium-independent alpha-latrotoxin receptor; CIRL; trkE; synapsin 1A;
KW epithelial discoidin domain receptor 1; synapsin 1B; neuroleptic;
KW tumour necrosis factor alpha; TNF-alpha; rat.
XX
OS Rattus sp.
XX
PN WO200175440-A2.
XX
PD 11-OCT-2001.
XX
PF 02-APR-2001; 2001WO-GB001486.
XX
PR 31-MAR-2000; 2000GB-00007880.
PR 26-MAY-2000; 2000GB-00012768.
XX
PA (WELF-) WELFIDE CORP.
XX
PI Cochran S, Paterson G, Ohashi Y, Morris B, Pratt J;
XX
DR WPI; 2002-010813/01.
DR N-PSDB; AAS16843.
XX
PT Novel chronic animal model of schizophrenia, useful for identifying anti-
PT psychotic drugs and genes that are associated with schizophrenia.
XX
PS Disclosure; Fig 8b; 79pp; English.
XX
CC The invention relates to YSG polynucleotide fragments for use in
CC diagnosing and/or developing treatments for schizophrenia using chronic
CC animal models. The polynucleotides and their encoded polypeptides are
CC used for identification of compounds which modulate the expression of YSG
CC molecules, leading to the manufacture of schizophrenia medicaments. The
CC sequences can also be used for testing candidate compounds for any effect
CC on the polypeptides. Anti-schizophrenic effects of a compound can be
CC determined by measuring local cerebral glucose utilisation (LCGU) or
CC comparing its expression level with that of a control group. The
CC sequences are useful in the identification of genes associated with
CC schizophrenic states and in the development of an antibody. The sequences
CC of the invention include phosphodiesterase 1-alpha, calcium-independent
CC alpha-latrotoxin receptors (CIRL)-1,2&3, epithelial discoidin domain
CC receptor 1 (trkE), netrin receptor (UNC5H1), synapsins 1A and AB and
CC tumour necrosis factor (TNF) alpha. This sequence represents rat netrin
CC receptor UNC5H1 (YSG7) polypeptide

XX

SQ Sequence 898 AA;

Query Match 96.8%; Score 4638; DB 5; Length 898;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 862; Conservative 17; Mismatches 19; Indels 0; Gaps 0;

Qy 1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
Db 1 MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy 61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSQQVEKVFGL 120
Db 61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSGLPTMEVRINVSQQVEKVFGL 120

Qy 121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
Db 121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy 181 VEWLRNEDLVDPSDLPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVIVY 240
Db 181 VEWLRNEDLVDPSDLPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSTSAAVIVY 240

Qy 241 VNNGWSTWTEWSVCSASCGRGWQKRSRSTNPAPLNGGAFCEGQNVQKTACATLCPVDGS 300
Db 241 VNNGWSTWTEWSVCSASCGRGWQKRSRSTNPAPLNGGAFCEGQNVQKTACATLCPVDGS 300

Qy 301 WSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTLDTRNCTS DL CVHSASGPEDVA 360
Db 301 WSSWSKWSACGLDCTHWRSRECSDPAPRNGGEECRGADLDTRNCTS DL CL HTASCPEDVA 360

Qy 361 LYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNP HLL 420
Db 361 LYIGLVAVAVCLFLLLALGLIYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNP HLL 420

Qy 421 TIQPDLSTTTTYQGSCLPRQDGSPKFQLTN GHLLSP LGGGRHTLHHSSPTSEAEEFVS 480
Db 421 TIQPDLSTTTTYQGSCLCSRQDGSPKFQLSNGHLLSP LGS GRHTLHHSSPTSEAEDFVS 480

Qy 481 RLSTQNYFRSLPRGTSNMTYGTNFNLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHK 540
Db 481 RLSTQNYFRSLPRGTSNMYGTNFNLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHK 540

Qy 541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHC GE PSPDSW SLRLKKQSCEGSW 600
Db 541 PEDVRLPLAGCQTLLSPVSCGPPGVLLTRPVILAMDHC GE PSPDSW SLRLKKQSCEGSW 600

Qy 601 EDVLHLGEEAPSHLYYCOLEASACYVFTEQLGRFALVGEALSVAAKRLKLLFAPVACT 660
Db 601 EDVLHLGEEPSHLYYCOLEAGACYVFTEQLGRFALVGEALSVAAKRLRLLFAPVACT 660

Qy 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHF KDSYHNLRLSIHDV PSSLW 720
Db 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHF KDSYHNLRLSIHDV PSSLW 720

Qy 721 KSKLLVSYQEIPFYHIWNGTQR YLHCTFTLERVSPSTS DLACKLWVWQVEGDGQSFSINF 780
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||

Db 721 KSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTSDLACKVWWQVEGDGQSFNINF 780
QY 781 NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKL 840
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 781 NITKDTRFAELLALESEGGVPALVGPSAFKIPFLIRQKIIASLDPPCSRGA DWRTLAQKL 840
QY 841 HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 898
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 841 HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC 898

RESULT 6

AAU97900

ID AAU97900 standard; protein; 898 AA.

XX

AC AAU97900;

XX

DT 27-AUG-2002 (first entry)

XX

DE Rat netrin binding membrane receptor UNC5H-1 protein.

XX

KW Netrin binding membrane receptor; receptor; UNC5H-1; Rat; nootropic;
KW neuroprotective; cytostatic; antiparkinsonian; cerebroprotective; cancer;
KW central nervous system; CNS; stroke; Parkinson's disease;
KW multiple sclerosis; Alzheimer's disease.

XX

OS Rattus sp.

XX

FH	Key	Location/Qualifiers
FT	Domain	152. .223
FT		/note= "Immunoglobulin domain "
FT	Domain	247. .294
FT		/note= "Thrombospondine type 1 domain "
FT	Domain	302. .348
FT		/note= "Thrombospondine type 1 domain"
FT	Region	361. .382
FT		/note= "Transmembrane region"
FT	Domain	495. .598
FT		/note= "ZU5 domain"
FT	Domain	817. .897
FT		/note= "Death domain"

XX

PN WO200233080-A2.

XX

PD 25-APR-2002.

XX

PF 15-OCT-2001; 2001WO-EP011891.

XX

PR 16-OCT-2000; 2000US-0240061P.

XX

PA (FARB) BAYER AG.

XX

PI Koehler RH;

XX

DR WPI; 2002-463314/49.

XX

PT Novel human netrin binding membrane receptor polypeptide and

PT polynucleotides for identifying modulating agents useful in treating
PT diseases e.g. Parkinson's disease, multiple sclerosis, stroke,
PT Alzheimer's disease.

PS Disclosure; Fig 3; 94pp; English.

xx

This invention relates to the DNA and protein sequences of a novel purified human netrin binding membrane receptor, UNC5H-1. The DNA sequence of the invention is useful as a probe for detecting a nucleic acid encoding the UNC5H-1 protein in a biological sample. The sequences of the invention are useful to screen for agents which decrease the activity of the UNC5H-1 protein. The sequences are also useful for screening agents which regulate (modulate) the activity of the protein of the invention. A pharmaceutical composition containing the protein of the invention or a reagent that modulates the activity of the UNC5H-1 protein may be useful for treating a UNC5H-1 dysfunction related disease such as cancer or a central nervous system (CNS) disorders (e.g., Parkinson's disease, multiple sclerosis, stroke and Alzheimer's disease). Fusion proteins comprising the UNC5H-1 protein are useful for generating antibodies and for in various assay systems, and the protein can be used as a bait protein in a two-hybrid assay or three-hybrid assay. The method of the invention is useful for detecting a coding sequence for the UNC5H-1 protein. The present sequence represents the Rat netrin binding membrane receptor UNC5H-1 protein of the invention.

xx

SQ Sequence 898 AA;

Db	361 LYIGLVAVAVCLFLLLALGLIYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPPLL	420
Qy	421 TIQPDLSTTTTYQGSLCPRQDGSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFVS	480
	: : : : : : : : : :	
Db	421 TIQPDLSTTTTYQGSLCSRQDGSPKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEDFVS	480
Qy	481 RLSTQNYFRSLPRGTSNM TYGTNFLLGRLMIPNTGISLLIPPD AIPRGKIYEIYLTLHK	540
	: : : : : : : : : :	
Db	481 RLSTQNYFRSLPRGTSNM MAYGTNFLLGRLMIPNTGISLLIPPD AIPRGKIYEIYLTLHK	540
Qy	541 PEDVRLPLAGCQTLLS PIVSCGPPGVLLTRPVILAMDHCGEPS PDSWLSRLKKQSCEGSW	600
	: : : : : : : : : :	
Db	541 PEDVRLPLAGCQTLLS PIVSCGPPGVLLTRPVILAMDHCGEPS PDSWLSRLKKQSCEGSW	600
Qy	601 EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALS VAAKRLKLLFAPVACT	660
	: : : : : : : : : :	
Db	601 EDVLHLGEEAPSHLYYCQLEAGACYVFTEQLGRFALVGEALS VAAKRLRLLFAPVACT	660
Qy	661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHF KDSYHNLRLSIHDVPSSLW	720
	: : : : : : : : : :	
Db	661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHF KDSYHNLRLSIHDVPSSLW	720
Qy	721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS DLA CKLWVWQVEGDGQSFSINF	780
	: : : : : : : : : :	
Db	721 KSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTS DLA CKVWWQVEGDGQSFNINF	780
Qy	781 NITKDTRFAELLALES EAGVPALVGPSAFKIPFLIRQKII SS LDPPCRRGADWRTLAQKL	840
	: : : : : : : : : :	
Db	781 NITKDTRFAELLALES EGGVPALVGPSAFKIPFLIRQKII ASLDPPCSRGADWRTLAQKL	840
Qy	841 HLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAA VAGLGQPDAGLFTVSEAE C	898
	: : : : : : : : : :	
Db	841 HLD SHLSFFASKPSPTAMILNLWEARHFPNGNLGQLAAA VAGLGQPDAGLFTVSEAE C	898

RESULT 7

AAM79128

ID AAM79128 standard; protein; 943 AA.

XX

AC AAM79128;

XX

DT 06-NOV-2001 (first entry)

XX

DE Human protein SEQ ID NO 1790.

XX

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.

XX

OS Homo sapiens.

XX

PN WO200157190-A2.

XX

PD 09-AUG-2001.

XX

PF 05-FEB-2001; 2001WO-US004098.

XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR N-PSDB; AAK52261.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
PS Claim 20; Page 4148-4150; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 943 AA;

Query Match 94.5%; Score 4526.5; DB 4; Length 943;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 863; Conservative 2; Mismatches 10; Indels 67; Gaps 4;

Qy 1 MAVRPGGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
| || | :|| | |||||||||||||||||||||||||||||||||||
Db 25 MTRRPSL-----MAGRQHGWSAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 76

Qy 61 VLLVCKAVPATQIFFKCNGEWRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVFGL 120
|||||||||||||||||||||||||||||||||||||||||||||
Db 77 VLLVCKAVPATQIFFKCNGEWRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVFGL 136

Qy 121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
||||||||||||||||| |||||||||||||||||||||||||||
Db 137 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 196

Qy 181 VEWLNRNEDLVDPSDLPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVIVY 240

Db	197	 VEWLRNEDLVDPSDLPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVIVY	256
Qy	241	-----VNGGWSTWTEWSVCSASCGRGWQKRSRSCTN	271
Db	257	 GGPRDSLVTGRGTAVPLGSDMWLSFSVRPVNGWSTWTEWSVCSASCGRGWQKRSRSCTN	316
Qy	272	PAPLNCGAFCEGQNQVQKTACATLCPVDGSPWSKWSACGLDCTHWRSRECSDPAPRNGG 	331
Db	317	PAPLNCGAFCEGQNQVQKTACATLCPVDGSPWSKWSACGLDCTHWRSRECSDPAPRNGG	376
Qy	332	EECQGTDLDTRNCTS DLCVH-----SASGPEDVALYVGLIAAVAVCLVLLL :	376
Db	377	EECQGTDLDTRNCTS DLCVHNSYTPAPTKAMLS PAAASGPEDVALYVGLIAAVAVCLVLLL	436
Qy	377	LVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNP HLLTIQPDLS TTTTYQGS 	436
Db	437	LVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNP HLLTIQPDLS TTTTYQGS	496
Qy	437	LCPRQDGSPPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAE FVSR LSTQNYFRSLPRGTS 	496
Db	497	LCPRQDGSPPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAE FVSR LSTQNYFRSLPRGTS	556
Qy	497	NMTYGTFNFLGGRLMIPNTGISL LIPPDAI PRGKIYEIYLTLHKPEDVRLPLAGCQTLLS 	556
Db	557	NMTYGTFNFLGGRLMIPNTGISL LIPPDAI PRGKIYEIYLTLHKPED-----	603
Qy	557	PIVSCGPPGVLLTRPVI LAMDHC GE PDSW SLRLKKQSCEGS WEDVLHLGEA PSHLYY 	616
Db	604	--VSCGPPGVLLTRPVI LAMDHC GE PDSW SLRLKKQSCEGS WEDVLHLGEA PSHLYY	661
Qy	617	CQLEASACYVFTEQLGRFALVGEALS VAAKRLKLLL FAPVACTSLEYNIRVYCLHDTHD 	676
Db	662	CQLEASACYVFTEQLGRFALVGEALS VAAKRLKLLL FAPVACTSLEYNIRVYCLHDTHD	721
Qy	677	ALKEVVQLEKQLGGQLIQEPRVLHF KDSYHNRLSIHDVSSLWKS KLLVSYQEIPFYHI 	736
Db	722	ALKEVVQLEKQLGGQLIQEPRVLHF KDSYHNRLSIHDVSSLWKS KLLVSYQEIPFYHI	781
Qy	737	WN GTQRYLHCTFTLERVSPSTS DLACKLWVWQVEGDGQSF SINF NITKDTRFAELLAES 	796
Db	782	WN GTQRYLHCTFTLERVSPSTS DLACKLWVWQVEGDGQSF SINF NITKDTRFAELLAES	841
Qy	797	EAGVPALVGPSAFKIPFLIRQKIISLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPT 	856
Db	842	EAGVPALVGPSAFKIPFLIRQKIISLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPT	901
Qy	857	AM I LNLWEARHFPNGNLS SQLAAA VAGLG QPDAGLFTVSEAEC 	898
Db	902	AM I LNLWEARHFPNGNLS QIAAA VAGLG QPDAGLFTVSEAEC	943

RESULT 8

AAU74818

ID AAU74818 standard; protein; 842 AA.

XX

AC AAU74818;

XX
DT 23-APR-2002 (first entry)
XX
DE Human REPTR 1 protein.
XX
KW REPTR; human; antiinflammatory; cytostatic; immunosuppressive; antiviral;
KW anti-HIV; antiarthritic; anticonvulsant; nootropic; neuroprotective;
KW antiallergic; antibody; immunogen; endometriosis;
KW gastrointestinal disorder; gastritis; oesophageal carcinoma;
KW Crohn's disease; irritable bowel syndrome; ulcerative colitis;
KW endocrine disorder; hypothalamus disorder; Kallman's disease;
KW autoimmune disease; inflammatory disease; infertility; receptor;
KW acquired immune deficiency syndrome; AIDS; rheumatoid arthritis; allergy;
KW osteoarthritis; diabetes mellitus; multiple sclerosis;
KW systemic lupus erythematosus; cell proliferative disorder; cancer;
KW developmental disorder; Duchenne muscular dystrophy;
KW Becker muscular dystrophy; neurological disorder; epilepsy;
KW Alzheimer's disease; Huntington's disease; reproductive disorder.
XX
OS Homo sapiens.
XX
PN WO200198354-A2.
XX
PD 27-DEC-2001.
XX
PF 21-JUN-2001; 2001WO-US019942.
XX
PR 21-JUN-2000; 2000US-0214027P.
PR 25-AUG-2000; 2000US-0228045P.
PR 12-DEC-2000; 2000US-0255104P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Griffin JA, Kallick DA, Tribouley CM, Yue H, Nguyen DB, Tang YT;
PI Lal P, Policky JL, Azimzai Y, Lu DAM, Graul R, Yao MG, Burford N;
PI Hafalia AJA, Baughn MR, Bandman O, Patterson C, Yang J, Xu Y;
PI Gandhi AR, Warren BA, Ding L, Sanjanwala MS, Duggan BM, Lu Y;
XX
DR WPI; 2002-090432/12.
DR N-PSDB; ABK15169.
XX
PT Twelve human receptors (referred to as REPTR-1 to REPTR-12), useful in
PT the diagnosis, treatment and prevention of gastrointestinal (e.g.
PT gastritis), autoimmune/inflammatory (e.g. osteoarthritis) and cell
PT proliferative (e.g. cancer) disorders.
XX
PS Claim 45; Page 111-113; 157pp; English.
XX
CC This invention relates to twelve human receptors cDNA sequences referred
CC to as REPTR-1 to REPTR-12), and the proteins encoded thereby. The
CC proteins of the invention may have antiinflammatory, cytostatic,
CC immunosuppressive, antiviral, anti-HIV, antiarthritic, muscular active
CC general, anticonvulsant, nootropic, neuroprotective, antiallergic
CC activities. The sequences of the invention may be used to produce REPTR
CC agonists or antagonists, and the protein sequences may be used to raise
CC anti-REPTR antibodies. These molecules and the REPTR polynucleotides and
CC polypeptides of the invention are useful in the diagnosis, treatment and

CC prevention of gastrointestinal (e.g. gastritis, oesophageal carcinoma,
CC Crohn's disease, irritable bowel syndrome, ulcerative colitis), endocrine
CC (e.g. hypothalamus disorder, Kallman's disease), autoimmune/ inflammatory
CC (e.g. acquired immune deficiency syndrome (AIDS), rheumatoid arthritis,
CC allergies, osteoarthritis, diabetes mellitus, multiple sclerosis,
CC systemic lupus erythematosus), cell proliferative (e.g. cancer),
CC developmental (e.g. Duchenne and Becker muscular dystrophy), neurological
CC (e.g. epilepsy, Alzheimer's disease, Huntington's disease) and
CC reproductive (e.g. infertility, endometriosis) disorders. Numerous other
CC examples of each disorder are given in the specification. The present
CC sequence represents the human REPTR1 protein sequence of the invention
XX
SQ Sequence 842 AA.

SQ Sequence 842 AA;

Query Match 92.1%; Score 4413; DB 5; Length 842;
Best Local Similarity 93.5%; Pred. No. 0;
Matches 840; Conservative 1; Mismatches 1; Indels 56; Gaps 1;

Qy 1 MAVRPGLWPALLGIVLAALWRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
||||| ||||| ||||| ||||| ||||| |||||

Qy 61 VLLVCKAVPATQIFFKCNGEWRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVFGL 120

Qy 121 EYWCQCVAWSSGTTSQKAYIRIARLRKNFEQEPLAKEVSLEQIVLPCRPPEGIPPAE 180
 |||||||
 Pb 121 EYWCQCVAWSSGTTSQKAYIRIARLRKNFEQEPLAKEVSLEQIVLPCRPPEGIPPAE 180

QY 241 VNGGWSTWTEWSVCSASCGRGWQKRSRSTNPAPLNGGAFCEGQNVQKTACATLCPVDGS 300
|||||

Db 245 WSPWSKWSACGLDCTHWRSCRECSDPAPRNGGEECQGTLDTRNCTS DLCVHTASGPEDVA 304

Qy 601 EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLL FAPVACT 660
||| ||| ||| ||| ||| ||| |||
Db 545 EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLL FAPVACT 604

Qy 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW 720
||| ||| ||| ||| ||| |||
Db 605 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW 664

Qy 721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS DLACKLWVWQVEGDGQSFSINF 780
||| ||| ||| ||| |||
Db 665 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS DLACKLWVWQVEGDGQSFSINF 724

Qy 781 NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDPPCRRGADWRTLAQKL 840
||| ||| ||| |||
Db 725 NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDPPCRRGADWRTLAQKL 784

Qy 841 HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 898
||| ||| ||| |||
Db 785 HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 842

RESULT 9

AAW78899

ID AAW78899 standard; protein; 556 AA.

XX

AC AAW78899;

XX

DT 25-MAR-2003 (revised)

DT 21-DEC-1998 (first entry)

XX

DE Human UNC-5 homologue UNC5H-1.

XX

KW UNC-5; UNC5H-1; human; netrin receptor; cell migration; axon guidance;
KW diagnosis; therapy.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
FT	Misc-difference	7 /note= "encoded by TG"
FT	Misc-difference	67 /note= "encoded by ATCT"
FT	Misc-difference	256 /note= "encoded by GC"
FT	Misc-difference	262 /note= "encoded by TG"
FT	Misc-difference	360 /note= "encoded by AG"
FT	Misc-difference	367 /note= "encoded by CC"
FT	Misc-difference	370 /note= "encoded by TC"
FT	Misc-difference	542 /note= "encoded by GG"

XX

PN WO9837085-A1.

XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-US003143.
XX
PR 19-FEB-1997; 97US-00808982.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Tessier-Lavigne M, Leonardo ED, Hinck L, Masu M, Keinomasu K;
XX
DR WPI; 1998-495364/42.
DR N-PSDB; AAW78899.
XX
PT Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and
PT the biopharmaceutical industry.
XX
PS Claim 1; Page 22-23; 32pp; English.
XX
CC UNC5H-1 and UNC5H-2 (see AAW78901) are human homologues of *Caenorhabditis elegans* UNC-5 protein. Their amino acid sequences were deduced from isolated unc5h cDNA clones (see AAV52941 and AAV52943) isolated from an embryonic brain cDNA library. The predicted proteins show similarity with UNC-5, possess 2 predicted Ig-like domains and 2 predicted thrombospondin type-1 repeats, a predicted membrane spanning region, and a large intracellular domain. They are predicted to be involved in cell migration and axon guidance, and are characterised as receptor proteins for netrins. Rat UNC5H-1 (see AAW78898) and UNC5H-2 (see AAW78900) proteins are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly from transfected host cells. The invention also provides unc-5 hybridisation probes and primers, vertebrate UNC-5-specific binding agents such as specific antibodies, and methods of making and using the subject compositions in diagnosis (e.g. genetic hybridisation screens for vertebrate unc-5 transcripts), therapy (e.g. gene therapy to modulate vertebrate unc-5 gene expression) and in the biopharmaceutical industry (e.g. as immunogens, reagents for modulating cell guidance, reagents for screening chemical libraries for lead pharmacological agents, etc.).
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 556 AA;

Query Match 58.8%; Score 2815; DB 2; Length 556;
Best Local Similarity 96.9%; Pred. No. 1.2e-225;
Matches 539; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

Qy 343 NCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 402
|||:|||| ||:|||||||||||||||||||||||||||||||||||||||||||||
Db 1 NCTSDLXVHTASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 60

Qy 403 GFQPVSINKPSKADNPILLTIQPDLSTTTTYQGSILCPQDGSPSPKFQLTNGHLLSPLGGG 462
|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 GFQPVSINKPSKADNPILLTIQPDLSTTTTYQGSILCPQDGSPSPKFQLTNGHLLSPLGGG 120

Qy 463 RHTLHHSSPTSEAEFFSRLSTQNYFRSLPRGTSNMTYGTNFNLGGRLMIPNTGISLLIP 522
|||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 RHTLHHSSPTSEAEFFSRLSTQNYFRSLPRGTSNMTYGTNFNLGGRLMIPNTGISLLIP 180

Qy	523 PDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHC	582
Db	181 PDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHC	240
Qy	583 SPDSWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALS	642
Db	241 SPDSWSLALKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALS	300
Qy	643 VAAAKRLKLLLFA PVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFK	702
Db	301 VAAAKRLKLLLFA PVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHIX	360
Qy	703 DSYHNLRLSIHDVPSSLWKS KL LV SYQEIPFYHIWNGTQR YLHCTFTLERVSPSTS	762
Db	361 DSYHNLXLSXHDVPSSLWKS KL LV SYQEIPFYHIWNGTQR YLHCTFTLERVSPSTS	420
Qy	763 KLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISS	822
Db	421 KLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISS	480
Qy	823 LDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAG	882
Db	481 LDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAG	540
Qy	883 LGQPDAGLFTVSEAEC 898	
Db	541 TPAGRWLLSQCSEAEC 556	

RESULT 10

AAB50691

ID AAB50691 standard; protein; 931 AA.

XX

AC AAB50691;

XX

DT 19-MAR-2001 (first entry)

XX

DE Human UNC5C protein SEQ ID NO:90.

XX

KW Human; Caenorhabditis elegans; UNC-5; splice variant; nematode worm;
KW protein-protein interaction; identification.

XX

OS Homo sapiens.

XX

PN WO200073328-A2.

XX

PD 07-DEC-2000.

XX

PF 02-JUN-2000; 2000WO-EP005108.

XX

PR 01-JUN-1999; 99GB-00012755.

XX

PA (DEVG-) DEVGEM NV.

XX

PI Van Criekinge W, Roelens I, Bogaert T, Verwaerde P;

XX

DR WPI; 2001-016508/02.

XX
PT Three variants of human unc-5C cDNAs (unc-5Cb, unc-5Cc and unc-5C8) and a
PT human unc-5HS1 cDNA, useful in yeast two hybrid experiments for
PT identifying unknown human cDNAs which encode proteins that interact with
PT the human unc-5C protein.

XX
PS Disclosure; Page 224-227; 246pp; English.
XX

CC The present invention describes 3 variants of human unc-5C cDNAs (unc-
CC 5Cb, unc-5Cc and unc-5C8) which correspond to alternatively spliced unc-
CC 5C transcripts, and a human unc-5HS1 cDNA which shares homology with the
CC Rattus norvegicus unc-5HS1 cDNA. Also described are assays based on
CC protein-protein-interactions between the unc-5 protein and a variety of
CC different interacting proteins. The unc-5C variant cDNAs and unc-5HS1
CC cDNA are useful in methods for identifying compounds which reduce or
CC inhibit the lethal phenotype associated with the expression of the unc-5
CC death domain in yeast. They are also useful in yeast two hybrid
CC experiments for identifying unknown human cDNAs which encode proteins
CC that interact with the human unc-5C protein. AAC90914 to AAC90971 and
CC AAB50646 to AAB50693 represent sequences used in the exemplification of
CC the present invention

XX
SQ Sequence 931 AA;

Query Match 57.5%; Score 2755; DB 4; Length 931;
Best Local Similarity 56.4%; Pred. No. 2.8e-220;
Matches 514; Conservative 154; Mismatches 215; Indels 28; Gaps 9;

Qy 9 PALLGIVLAAWLRGSGAQQS---ATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC 65
||| :|: | ||| : | : |: |||||:|||: ||||||| | |
Db 26 PAL--ALLSASGTGSAAQDDDFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYC 83

Qy 66 KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVRQQVEKVGLEEWQC 125
|| |||||:|||| | ||| | ||::: | :||| || | :|||||:||| |:|||
Db 84 KASPAHQIYFKCNSEWVHQKDHVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQ 143

Qy 126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWL 185
|| |||||:|||||:|||:||| ||| ||||||| ||||||| ::| ||||||| |||||||:
Db 144 CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK 203

Qy 186 NEDLVDPSDLPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVYVNGGW 245
|||::||| | | ||| :|||::|||:||| ||||||| |||||:||| :| | ||||||| |||
Db 204 NEDIIDPVEDRFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW 263

Qy 246 STWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWPWS 305
||| ||| :||||| |||: |||:|||::| | | :||| ||| :| :||| |||:
Db 264 STWTEWSVCNSRCGRGYQKRTCTNPAPLNGGAFCEGQSVQKIACTTLCVPDGRWTPWS 323

Qy 306 KWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTSMLCVHSASGPEDVALYVGL 365
||| ||| :||||| |||: |||:|||::| | | :||| ||| :| :||| |||:
Db 324 KWSTCGTECTHWRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQTAPDSDDVALYVGI 383

Qy 366 -IAAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPPLLTIQP 424
||| ||| : ::| : ||| : | :||| | | |||||:||| ::| | | :|
Db 384 VIAVIVCLAISVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVPP 440

Qy 425 DLSTTTTYQGSLCPRQDGSPKFQLTNGHLLSPLGGGRHTLHHS---SPTSEAEFFVS 480

Db	:: : : : : : :::: :	
	441 DLTSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNTSGAVSPQDDLSEFTS	499
Qy	481 RLS---TQNYF-----RSLPRGT--SNMTYGTNFNFLGGRLMIPNTGISSLIPPAI	526
	: : : : : :	
Db	500 KLSPQMTQSLLNEALSLKNQSLARQTDPSTAFGSFNSLGGHLIVPNSGVSLIPAGAI	559
Qy	527 PRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVLAMDHCGEPSD	586
	: :: : : : : : : : : :	
Db	560 PQGRVYEMYVTVHRKETMRPPMDDSQTLTPVVSCGPPGALLTRPVVLTMHHCADPNTED	619
Qy	587 WSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAA	646
	: : : : : : : :	
Db	620 WKILLKNQAAQGQWEDVVVGEENFTTPCYIKLDAEACHILLENLSTYALVGHSTTKAAA	679
Qy	647 KRLKLLLFAFPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYH	706
	: : : : : :	
Db	680 KRLKLAIGPLCCSSLEYSIRVYCLDDTQDALKEILHLERQTGGQLLEEPKALHFKGSTH	739
Qy	707 NLRLSIHDVPSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSSDLACKLWV	766
	: : : : :	
Db	740 NLRLSIHDIAHSLWKSLLAKYQEIPFYHVWSGSQRNLHCTFTLERFSLNTVELVCKLCV	799
Qy	767 WQVEGDGQSFSINFNITKDTRFAELLALESAGVPALVGPSAFKIPFLIRQKIISSLDP	826
	:: : : : : : : : : : : : : : : :	
Db	800 RQVEGEGQIFQLNCTVSEEPTGIDLPLLDPANTITTGTGPSAFSIPPLPIRKLCSSLDAP	859
Qy	827 CRRGADWRTLAQKLHLDHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAGLGQP	886
	: : : : : : : : : : : : :	
Db	860 QTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDWEAQNFPDGNLSMLAAVLEEMGRH	919
Qy	887 DAGLFTVSEAE	897
	: : : :	
Db	920 ETVVSLAAEQQ	930

RESULT 11

ADE63098

ID ADE63098 standard; protein; 931 AA.

XX

AC ADE63098;

XX

DT 29-JAN-2004 (first entry)

XX

DE Human Protein AAC67491, SEQ ID NO 9033.

XX

KW Human; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; chronic constriction injury; CCI;

KW spared nerve injury; SNI; Chung.

XX

OS Homo sapiens.

XX

PN WO2003016475-A2.

XX

PD 27-FEB-2003.

XX

PF 14-AUG-2002; 2002WO-US025765.

XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; AAC67491.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 931 AA;

Query Match 57.5%; Score 2755; DB 7; Length 931;
Best Local Similarity 56.4%; Pred. No. 2.8e-220;
Matches 514; Conservative 154; Mismatches 215; Indels 28; Gaps 9;

QY 9 PALLGIVLAAWLRGSGAQQS---ATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLVLC 65
||| :|:|| |||| : | |:|||||:|||: |||||||| |||
Db 26 PAL--ALLSASGTGSAAQDDDFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYC 83
QY 66 KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSQQVEKVFGLEEYWCQ 125

Db	84	KASPATQIYFKCNSEVHQQDHIVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQ	143
Qy	126	CAWSSSGTTKSQAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLR	185
Db	144	CAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK	203
Qy	186	NEDLVDPSDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVIVVNGGW	245
Db	204	NEDIIDPVEDRNFYITIDHNLIQARLSDTANYTCVAKNIVAKRKSTTATVIVIVVNGGW	263
Qy	246	STWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWS	305
Db	264	STWTEWSVCNSRCGRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIACTTLCVPDGRWTPWS	323
Qy	306	KWSACGLDCTHWRSCSDPAPRNGGEECQGTDLDRNCTS DLCVHSASGPEDVALYVGL	365
Db	324	KWSTCGTECTHWRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQTAPDSDDVALYVGI	383
Qy	366	-IAVAVCLVLLLLVILVYCRKKEGLSDVADSSI LTSGFQPVSIKPSKADNPHLLTIQP	424
Db	384	VIAVIVCLAISVVVALFVYRKNRDFESDIIDSSALNGGFQPVNKAARQD---LLAVPP	440
Qy	425	DLSTTTTYQGS LCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHS----SPTSEAEFFVS	480
Db	441	DLTSAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNTSGAVSPQDDLSEFTS	499
Qy	481	RLS---TQNYF-----RSLPRGT--SNMTYGTNFLLGRLMIPNTGISLLIPPDAI	526
Db	500	KLSPQMTQSLLNEA SLKNQSLARQTDPSC TAFGSFNLSGGHLIVPNNSGVSLIPAGAI	559
Qy	527	PRGKIYEYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHC GEPS PDS	586
Db	560	PQGRVYEMYVTVHRKETMRPPMDDSQTLLTPV VSCGPPGALLTRPVVLTMHHCADPNTED	619
Qy	587	WSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAA	646
Db	620	WKILLKNQAAQGQWEDVVVVGEEENFTTPCYIKLDAEACHILTENLSTYALVGHSTTKAAA	679
Qy	647	KRLKLLL FAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYH	706
Db	680	KRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEILHLERQTGGQLLEEPKALHFKGSTH	739
Qy	707	NLRLSIHDVPSLWKS KLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS DLACKLWV	766
Db	740	NLRLSIHDIAHSLWKS KLLAKYQEIPFYHVWSGSQRNLHCTFTLERFSLNTVELVCKLCV	799
Qy	767	WQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAVGPSAFKIPFLIRQKIISSL DPP	826
Db	800	RQVEGEGQIFQLNCTVSEEPTGIDLP LLDPANTITTVTGPSAFSIPLPIRQKLCSSL DAP	859
Qy	827	CRRGADWRTL A QKLHLD SHLSFFASKPSPTAMI LN IWEARHF PNGNLSQ LAAVAGLGQP	886
Db	860	QTRGH DWMLAHKLNLD RYLNYFATKSSPTGVILD LWEAQNF PDGNLSMLAAVLEEMGRH	919
Qy	887	DAGLFTVSEAE 897	

RESULT 12
ABG11551
ID ABG11551 standard; protein; 982 AA.
XX
AC ABG11551;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #11542.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS75738.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 41910; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 982 AA;

Query Match 57.5%; Score 2755; DB 4; Length 982;
Best Local Similarity 56.4%; Pred. No. 3e-220;
Matches 514; Conservative 154; Mismatches 215; Indels 28; Gaps 9;

Qy 9 PALLGIVLAAWLRGSGAQQS---ATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLVC 65
Db 77 PAL--ALLSASGTGSAAQDDDFHELPETFPSPDPPEPLPHFLIEPEEAYIVKNKPVNLYC 134

Qy 66 KAVPATQIFFKCNGEWRQVDHIERSTDGSSGLPTMEVRINVSRRQQVEKVFGLEEYWCQ 125
Db 135 KASPAHQIYFKCNSEWVHQKDHVDERVDETSGLIVREVSIEISRQQVEELFGPEDIWCQ 194

Qy 126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVWLR 185
Db 195 CVAWSSAGTTKSRSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVWLK 254

Qy 186 NEDLVDPSDLPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVIVYVNGGW 245
Db 255 NEDIIDPVEDRNFYITIDHNLIQKARLSDTANYTCVAKNIVAKRSTTATVIVYVNGGW 314

Qy 246 STWTEWSVCASCGRGWQKRSRSCTNPAPLNGGAFCEGQNQKTAATLCPVDGSWSPWS 305
Db 315 STWTEWSVCNSRCGRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIACTTLCVPDGRWTPWS 374

Qy 306 KWSACGLDCTHWRSRECSDPAPRNGGECCQGTDLDTRNCTSDLCVHSASGPEDVALYVGL 365
Db 375 KWSTCGTECHWRRRECTAPAKNGGKDCDGLVLQSKNCTDGLCMQTAPDSDDVALYVGI 434

Qy 366 -IAAVAVCLVLLLLVLIIVYCRKKEGLDSVDASSILTSGFQPVSIKPSKADNPHLTIQP 424
Db 435 VIAVIVCLAISVVVALFVYRKNHDFESDIIDSSALNGGFQPVNIKAARQD---LLAVPP 491

Qy 425 DLSTTTTYQGSCLPRQDGSPKFQLTNGHLLSPLGGRHTLHHS---SPTSEAEFVS 480
Db 492 DLTSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNTSGAVSPQDDLSEFTS 550

Qy 481 RLS---TQNYF-----RSLPRGT--SNMTYGTNFLLGGRLMIPNTGISLLIPPDAI 526
Db 551 KLSPQMTQSLLENEALSLKNQSLARQTDPSCTAFGSFNSLGGHLIVPNSGVSLIPAGAI 610

Qy 527 PRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSD 586
Db 611 PQGRVYEMYVTVHRKETMRPPMDDSQTLLTPVSCGPPGALLTRPVVLTMHCADPNTED 670

Qy 587 WSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAA 646
Db 671 WKILLKNQAAQGQWEDVVVGEENFTTPCYIKLDAEACHILLENLSTYALVGHSTTKAAA 730

Qy 647 KRLKLLLFAFPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYH 706

|| || || : | | : | : || || || | | || || || : | | : | | || || | |
Db 731 KRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEILHLERQTGGQLLEEPKALHFKGSTH 790

Qy 707 NLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS LACKLWV 766
| | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 791 NLRLSIHDIAHSLWKS KLLAKYQEIPFYHVWSGSQRNLHCTFTLERFS LNTVELVCKLCV 850

Qy 767 WQVEGDGQSFSINFNITKDTRFAELLALESAGVPALVGPSAFKIPFLIRQKIISSLDPP 826
| | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 851 RQVEGEQEIFQLNCTVSEEPTGIDPLPLDPANTITTVTGPSAFS IPLPIRQKLCSSLDAP 910

Qy 827 CRRGADWR TLAQKLHLD SHLSFFASKPSPTAMILNLWEARHF PNGNLSQLAAVAGLGQP 886
| | | | | | : | | | | | | : | | | | | | : | | | | | | | | | | | | : | :
Db 911 QTRGH DWRMLAHKL NLD RYLN YFATKSSPTGVILD LWEAQNF PDGNLSMLAAVLEEMGRH 970

Qy 887 DAGLFTVSEAE 897
: : : | :
Db 971 ETVVSLAAEGQ 981

RESULT 13

ADE63096

ID ADE63096 standard; protein; 945 AA.

XX

AC ADE63096;

XX

DT 29-JAN-2004 (first entry)

XX

DE Rat Protein AAB57679, SEQ ID NO 9031.

XX

KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX

OS Rattus norvegicus.

XX

PN WO2003016475-A2.

XX

PD 27-FEB-2003.

XX

PF 14-AUG-2002; 2002WO-US025765.

XX

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX

PA (GEHO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX

PI Woolf C, D'urso D, Befort K, Costigan M;

XX

DR WPI; 2003-268312/26.

DR GENBANK; AAB57679.

XX

PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.

XX

PS Claim 1; Page; 1017pp; English.

xx

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a rat protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp://wipo.int/pub/published_pct_sequences.

xx

SQ Sequence 945 AA;

Query Match 53.8%; Score 2578.5; DB 7; Length 945;
Best Local Similarity 53.0%; Pred. No. 1.5e-205;
Matches 509; Conservative 142; Mismatches 231; Indels 79; Gaps 17;

Qy

1 MAVRPGLWPALLGIVLAAW----LRG--SGAQQSATVANPVPGANPDLLPHFLVEPEDV 53

Db

1 MRARSGARGALLALLCWDPTPSLAGIDSGGQ---ALPDSEPSAPAEOLPHFELLPEADA 57

Ov

54 YIVKKNKPVLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSQQV 113
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Ph

58 YIVKNKPVELHCBRAEPAOTIYEKCNGEWYSOKGHVTESLDEATGLRIBEVOLAEVSROOV 117

QV

Dh

118 EELEGLEDYWCOCVVAWSSSGTTKSBRAYIBTAYI-BKNEOFLAKEVPLDHENVYLLOCBPP 177

QV

174 EGI PPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSA 233
||: : ||:: ||: : ||: | : | . : | : |: : |: : |: |: |

Dh

178 EGYRVAFEVWIKNEDVIRAOOTNEILLTIDHNLTIBOARISTDANXTCVAKNTIVAKRPSL 237

QV

234 SAAIVYVNGGWSTWTEWSVCASACGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACAT 293

Pb

238 TATVLYVYNGCWS SWAEWSRGNSRGCRCWOKRTPTCTNPARLNGCAECEGCONCOKTACTT 237

Q3

294 LCPVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEQGTDLDTRNCTSDLCV--- 350

Db 298 VCPVDGAWTEWSKWSACSTECAHWRRECMAAPPQNGGRDCSGTLLDSKNCTDGLCVLNQ 357
 ,
 QY 351 HSASGPE-----DVALYVGL-IAAVAVCLVLLLVLVILVYCRKKEGLDSDVADSS-IL 400
 : : |: | | | | : || | | : | : || | | : |: || | |
 Db 358 RTLNDPKSRPLEPSGDVALYAGLVAVFVVLAVLMAVGIVYRRNCRDFDTDITDSSAAL 417
 QY 401 TSGFQPVSISKPSKADNPHEL--TIQPDLSSTTTTYQGSCLCPRQDGSPKFQLTNGHLLSP 458
 | | | |: | : | | | : | |: : | | : | | : | : || | | | |
 Db 418 TGGFHPVNFKTARPSNPQLLHPSAPPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLLDP 476
 QY 459 L-----GGG-----RHTLHHSSPTSEAEFFVS 480
 | | | | | | | : :
 Db 477 LPSLKIKVYDSTSSTIGSGAGLADGADLLGVLPPTYPGDFSRDTHFLHLRS-----A 527
 QY 481 RLSTQNYFRSLPRGTSNMTRYGTFNFLGGRMLIPNTGISLILLPPDAIPRGKIYEIYLTLHK 540
 | : | : | | | | : | | | | | | | | : | | | | | | | | : | | : |
 Db 528 SLGSQ-HLLGLPRDPSSSVSGTFGCLGGRLTIPGTGVSLVNGAIPQGKFYDLYLRINK 586
 QY 541 PEDVRLPLA-GCQTLLSPLIVSCGPPGVLLTRPVLAMDHCGEPSWLSRLKKQSCEGS 599
 | | | | : | | | | : | | | | | | | | | | | | | | | : | | : |
 Db 587 TEST-LPLSEGSQTVLSPSVTCGPTGLLCRPVVLTVPHCAEVIAQDWIFQILKTQAHQGH 645
 QY 600 WEDVHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLFAPVAC 659
 | | : | | | : | | | | | | | | | | | | | | | | | | | | | | |
 Db 646 WEEVVTLDDEETLNTPCYCQLEAKSCHILLDQLGYVFTGESYSRSAVKRLQLAIFAPALC 705
 QY 660 TSLEYNIRVYCLHDTHDALKEVQVLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVSSL 719
 | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 706 TSLEYSLRVYCLEDTPAALKEVLELERTLGGYLVEEPKTLLFKDSYHNLRLSLHDIPHAB 765
 QY 720 WKSKLLVSYQEIPFYHIWNGTQRYLHCTTLERVSPSTSDDLACKLWWQVEGDGQSFSIN 779
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 766 WRSKLLAKYQEIPFYHWNGSQKALHCTTLERHSLASTEFTCKVCVRQVEGEGQIFQLH 825
 QY 780 FNITKDTRFAELLALESEAGVPAL--VGPSAFKIPFLIRQKIISSLPPCRRGADWRTLA 837
 : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 826 TTLA-ETPAGSLDALCSAPGNAATTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRLLA 884
 QY 838 QKLHLDSHLSFFASKPSPTAMIQLNWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 897
 | | | : | | : | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 885 QKLSMDRYLNYFATKASPTGVILDWEARQQDDGDLNSLASALEEMGKSEMLVAMTTDGD 944
 QY 898 C 898
 |
 Db 945 C 945

RESULT 14
 AAW78900
 ID AAW78900 standard; protein; 943 AA.
 XX
 AC AAW78900;
 XX
 DT 25-MAR-2003 (revised)
 DT 21-DEC-1998 (first entry)
 XX
 DE Rat UNC-5 homologue UNC5H-2.

XX
KW UNC-5; UNC5H-2; rat; netrin receptor; cell migration; axon guidance;
KW diagnosis; therapy.
XX
OS Rattus sp.
XX
FH Key Location/Qualifiers
FT Peptide 148. .161
FT /note= "peptide used to raise rabbit polyclonal antisera"
FT Misc-difference 753
FT /note= "encoded by CG"
FT Peptide 909. .924
FT /note= "peptide used to raise rabbit polyclonal antisera"
XX
PN WO9837085-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-US003143.
XX
PR 19-FEB-1997; 97US-00808982.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Tessier-Lavigne M, Leonardo ED, Hinck L, Masu M, Keinomasu K;
XX
DR WPI; 1998-495364/42.
DR N-PSDB; AAV52942.
XX
PT Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and
PT the biopharmaceutical industry.
XX
PS Claim 1; Page 24-26; 32pp; English.
XX
CC UNC5H-1 and UNC5H-2 (see AAW78900) are rat homologues of *Caenorhabditis*
CC elegans UNC-5 protein. Their amino acid sequences were deduced from
CC isolated unc5h cDNA clones (see AAV52940 and AAV52942) isolated from an
CC E18 brain cDNA library. The predicted proteins show similarity with UNC-
CC 5, possess 2 predicted Ig-like domains and 2 predicted thrombospondin
CC type-1 repeats, a predicted membrane spanning region, and a large
CC intracellular domain. They are predicted to be involved in cell migration
CC and axon guidance, and are characterised as receptor proteins for
CC netrins. Human UNC5H-1 (see AAW78899) and UNC5H-2 (see AAW78901) proteins
CC are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly
CC from transfected host cells. The invention also provides unc-5
CC hybridisation probes and primers, vertebrate UNC-5-specific binding
CC agents such as specific antibodies, and methods of making and using the
CC subject compositions in diagnosis (e.g. genetic hybridisation screens for
CC vertebrate unc-5 transcripts), therapy (e.g. gene therapy to modulate
CC vertebrate unc-5 gene expression) and in the biopharmaceutical industry
(e.g. as immunogens, reagents for modulating cell guidance, reagents for
CC screening chemical libraries for lead pharmacological agents, etc.).
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 943 AA;

Query Match

53.7%; Score 2571.5; DB 2; Length 943;

Best Local Similarity 53.3%; Pred. No. 5.8e-205;
Matches 504; Conservative 142; Mismatches 221; Indels 79; Gaps 16;

Qy 9 PALLGIVLAAWLRGSGAQQSATVANPVGANPDLLPHFLVEPEDVYIVKNKPVLVCKAV 68
Db 21 PSLAGI-----DSGAQ---GLPDSFPSAPAEGQLPHFLLEPEDAYIVKNKPVELHCRAF 70

Qy 69 PATQIFFKCNGEWRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVGLEEYWCQCVA 128
Db 71 PATQIYFKCNGEWSQKGHVTQESLDEATGLRIREVQIEVSRRQQVEELFGLEDYWCQCVA 130

Qy 129 WSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLKNED 188
Db 131 WSSSGTTKSRRRAYIRIAYLRKNFDQEPLAKEVPLDHEVLLQCRPPEGVPVAEVEWLKNED 190

Qy 189 LVDPSDLPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAASAVIVYVNGGWSTW 248
Db 191 VIDPAQDTNFLLTIDHNLIIQARLSDTANYTCVAKNIVAKRRSTTATVIVYVNGGWSSW 250

Qy 249 TEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPSKWS 308
Db 251 AEWSPCSNCGRGWQKRTRTCTNPAPLNGGAFCEGQACQKTAJTVCPVDAWTEWSKWS 310

Qy 309 ACGLDCTHWRSRECSDPAPRNGEECQGTDDLDRNCTS DLCV--HSASGP----- 357
Db 311 ACSTECAHWRSRECMAPPQNGGRDCSGTLLDSKNCTDGLCVLNQRTLNDPKSRPLEPSG 370

Qy 358 DVALYVGL-IAVAVCLVLLLVLILVYCRKKEGLDSVDSS-ILTSGFQPVSIKPSKAD 415
Db 371 DVALYAGLVVAVFVVLAVALMAVGIVYRRNCRDFDTIDSSAALTGGFHPVNFKTARPS 430

Qy 416 NPHLL--TIQPDLSTTTTYQGSILCPRQDGSPPKFQLTNGHLLSPL----- 459
Db 431 NPQLLHPSAPPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLLDPLPSLKIKVYDSSTIG 489

Qy 460 -GGG-----RHTLHHSSPTSEAEFVSRSLSTONYFRSLPRGT 495
Db 490 SGAGLADGADLLGVLPNGTYPGDFSRDTHFLHLRS-----ASLGSQ-HLLGLPRDP 539

Qy 496 SNMTYGTFNFLGGRLMIPNTGISLLIPDAIPRGKIYEIYLTLHKPEDVRLPLA-GCQTL 554
Db 540 SSSVSGTFCLGGRLTIPGTGVSLVPNGAIPOGKFYDLYLRINKTEST-LPLSEGSQTV 598

Qy 555 LSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSWEDVLHLGEEAPSHL 614
Db 599 LSPSVTCGPTGLLLCRVVLTVPHCAEVIAGDWIFQLKTQAHQGHWEVVTLDEETLNTP 658

Qy 615 YYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLFAPVACTSLEYNIRVYCLHDT 674
Db 659 CYCQLEAKSCHILLDQLGTYVFTGESYSRSAVKRLQLAIFAPALCTSLEYSLRVYCLEDT 718

Qy 675 HDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSLWKS KLLVSYQEIPFY 734
Db 719 PAALKEVLELERTLGGYLVEEPKTLFFKDSYHNLRLSLHDIPHAWRSKLLAKYQEIPFY 778

Qy 735 HIWNGTQRYLHCTFTLERVSPSTS D LACKLWVWQVEGDGQSFSINFNITKDTRFAELLAL 794
|:||||:|: |||||||| | :::: | :| | ||||:| | :: : | :| | ||

Db 779 HVWNGSQKALHCTFTLERHSLASTEFTCKVCVRQVEGEGQIFQLHTTLA-ETPAGSLDAL 837
QY 795 ESEAGVPAL--VGPSAFKIPFLIRQKIISSLDPCCRRGADWRTLQKLHLDSHLSFFASK 852
| | | :|| |||| | |||| :||| | || ||| ||||| :| :||::||:|
Db 838 CSAPGNAATTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRLLAQKLSDRYLNYFATK 897
QY 853 PSPTAMILNLWEARHFPGNQLSQLAAAVAGLGQPDAGLFTVSEAEC 898
||| :||:||||| :||:|: |||: :||: : : :: :|
Db 898 ASPTGVILDLWEARQQDDGDLNSLASALEEMGKSEMLVAMTTDGDC 943

RESULT 15

AAO18734

ID AAO18734 standard; protein; 933 AA.

XX

AC AAO18734;

XX

DT 24-OCT-2002 (first entry)

XX

DE Human NOV1a protein.

XX

KW Human; NOVX; autoimmune disease; cancer; infection; inflammatory disease;
KW storage disorder; muscle disorder; neurodegenerative disorder; nootropic;
KW developmental defect; neuroprotective; antiparkinsonian; hypotensive;
KW hypertensive; haemostatic; cardiant; antianginal; dermatological;
KW immunosuppressive; antiinflammatory; virucide; antibacterial; anti-HIV;
KW antiparasitic; antiallergic; antiasthmatic; antirheumatic; antiarthritic;
KW vulnerability; anorectics; antidiabetic; immunomodulator; antipsoriatic;
KW nephrostatic; kerolytic; antiulcer; cerebroprotective; anticonvulsant;
KW antiinfertility; antimanic; antidepressant; metabolic; cytostatic;
KW tranquilizer; analgesic.

XX

OS Homo sapiens.

XX

PN WO200257450-A2.

XX

PD 25-JUL-2002.

XX

PF 29-NOV-2001; 2001WO-US048922.

XX

PR 29-NOV-2000; 2000US-0253834P.

PR 30-NOV-2000; 2000US-0250926P.

PR 25-JAN-2001; 2001US-0264180P.

PR 20-AUG-2001; 2001US-0313656P.

PR 05-OCT-2001; 2001US-0327456P.

PR 28-NOV-2001; 2001US-00327456.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Edinger S, Macdougall JR, Millet I, Ellerman K, Stone DJ;

PI Gerlach V, Grosse WM, Alsobrook JP, Lepley DM, Rieger D, Burgess CE;

PI Casman SJ, Spytek KA, Boldog FL, Li L, Padigaru M, Mishra V;

PI Paturajan M, Shenoy S, Rastelli L, Tchernev VT, Vernet CAM;

PI Zerhusen BD, Malyankar UM, Guo X, Miller CE, Gangolli EA;

XX

DR WPI; 2002-590741/63.

DR N-PSDB; ABT06279.

xx

PT Novel isolated polypeptide, designated NOVX, useful for treating or preventing in NOVX-associated disorders e.g. cardiomyopathy, atherosclerosis, diabetes, cancer, allergy, asthma, Crohn's disease.

xx

PS Claim 1; Page 13; 353pp; English.

xx

CC The present invention provides the protein and coding sequences of
CC several novel human proteins, designated NOVX. These can be used in the
CC treatment of, amongst others, cancers, autoimmune diseases, infections,
CC inflammatory diseases, storage disorders, muscle disorders,
CC neurodegenerative diseases and developmental defects. The present
CC sequence is a protein of the invention

xx

SQ Sequence 933 AA;

Query Match 53.5%; Score 2563.5; DB 5; Length 933;
Best Local Similarity 53.4%; Pred. No. 2.7e-204;
Matches 501; Conservative 147; Mismatches 245; Indels 45; Gaps 15;

QY 1 MAVRPGILWPALLGIVIAAW--LRGSGAQQ-SATVANPVPGANPDLLPHFLVEPEDVYIV 56
| | | ||| :| | | :| | : :| | | :||:|| | :| | | | | |

Qy 57 KNKPVLVCKAVPATQIFFKCNGEWVRQVDHIERSTDGSSGLPTMEVRINVSRRQQVEKV 116
| | | | : | : | | | | : | | | : | : | : | | | | | | : | : | | | | | | | | : |
Pf 61 KNNKPVLVCKAVPATQIFFKCNGEWVRQVDHIERSTDGSSGLPTMEVRINVSRRQQVEKV 120

Qy 117 FGLEEWYCQCVAWSGGTTSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGI 176
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||| :||| :||| :|||:
E 121

Qy 177 PPAEVEWLRNEDLVDPSLDPNVYITREHSLVRQARLADTANYTCVAKNIVARRSASAA 236
| :||||||:||||:|||: | | :| :|||:|||||:|||||:|||||:||| :|

Qy 237 VIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSTNPAPLNGGAFCEGQNQKTACATLCP 296
|||:|||||:| ||| ||| |||:|||:|||:|||||:|||:|||:|||

Qy 297 VDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTLDLDRNCTSDLCVH-SASG 355
||:||:|: ||||||| :| ||||||| | :||| :| || |||:||| ||: |||

Qy 356 PEDVALYVGL-IAVAVCLVLLLVLILVYCRKKEGLDSDVADSS-ILTSGFQPVSIKPSK 413
| ||| || :| :| :| :| :|| | :| :| ||| || || ||| :| ::|

Qy 414 ADNPHLL--TIQPDLS||||TYQGS||CPRQDGSPKFQLTNGHLLSPLGGGRHTLHHSSP 471
|| || :: ||| : : || | : || : | : || || || | : :: ||

Qy 472 T-----SEAEVFVSRLSTQNY-----FRS-----LPRGTSNMTYGT F 503
| :: : | | || | : |||

Qv 504 NEICCRIMIPNTGSIITIRPDAIRPKIVELVITIHKPFDVPLIA-CCOTIISPRIVSCC 562

Db	538	GCLGGRLSIPGTGVSLVPNGAIPQGKFYEMYLLINKAEST-LPLSEGTQTVLSPSVTCG	596
Qy	563	PPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEAS	622
Db	597	PTGLLLCRPVILTMPHCAEVSA RDWIFQLKTQAHQGHWEVVTLDEETLNTPCYCQLEPR	656
Qy	623	ACYVFTEQLGRFALVGEALSVAAKRLKLLFAPVACTSLEYNIRVYCLHDTHDALKEVV	682
Db	657	ACHILLDQLGTYVFTGESYSRSAVKRLQLAVFAPALCTSLEYSLRVYCLEDTPVALKEVL	716
Qy	683	QLEKQLGGQLIQEPRVLFHFKDSYHNLRSLIHDVPSLWKSKLIVSQEIPFYHIWNGTQR	742
Db	717	ELERTLGGYLVEEPKPLMFKD SYHNRLSLHDLPHAHWRSKLLAKYQEIPFYHIWSGSQK	776
Qy	743	YLHCTFTLERVSPSTS D LACKLWVWQVEGDQGSFSINFNITKDTRFAELIALESEAG--V	800
Db	777	ALHCTFTLERHSLA STELTCKICVRQVEGEGQI FQLHTLA-ETPAGSLDTLCSAPGSTV	835
Qy	801	PALVGPSAFKIPFLIRQKIISLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMIL	860
Db	836	TTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRMLAQKLSMDRYLN YFATKASPTGVIL	895
Qy	861	NLWEARHFPGNQLSQA AAVAGLGQPDAGLFTVSEAEC	898
Db	896	DLWEALQDDGDLNSLASALEEMGKSEMLVAVATDGDC	933

Search completed: July 6, 2004, 14:33:27
Job time : 68 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2004, 14:32:16 ; Search time 23 Seconds
(without alignments)
2015.657 Million cell updates/sec

Title: US-10-624-932-2

Perfect score: 4791

Sequence: 1 MAVRPGLWPALLGIVLAAWL.....AVAGLGQPDAGLFTVSEAEC 898

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*

5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query					Description
No.	Score	Match	Length	DB	ID	
1	4638	96.8	898	2	US-08-808-982-5	Sequence 5, Appli
2	4638	96.8	898	3	US-09-306-902A-5	Sequence 5, Appli
3	2815.5	58.8	557	2	US-08-808-982-6	Sequence 6, Appli
4	2815.5	58.8	557	3	US-09-306-902A-6	Sequence 6, Appli
5	2571.5	53.7	943	2	US-08-808-982-7	Sequence 7, Appli
6	2571.5	53.7	943	3	US-09-306-902A-7	Sequence 7, Appli
7	296.5	6.2	1172	1	US-08-313-288B-19	Sequence 19, Appli
8	294	6.1	102	2	US-08-808-982-8	Sequence 8, Appli
9	294	6.1	102	3	US-09-306-902A-8	Sequence 8, Appli
10	268.5	5.6	239	5	PCT-US93-01652-1	Sequence 1, Appli
11	268.5	5.6	1170	1	US-08-313-288B-20	Sequence 20, Appli

12	249.5	5.2	441	3	US-08-985-526-3	Sequence 3, Appli
13	243	5.1	469	1	US-08-313-288B-15	Sequence 15, Appli
14	238	5.0	218	3	US-08-985-526-1	Sequence 1, Appli
15	234.5	4.9	1395	3	US-09-540-245A-15	Sequence 15, Appli
16	215	4.5	788	2	US-08-918-914-4	Sequence 4, Appli
17	210	4.4	1651	3	US-09-540-245A-18	Sequence 18, Appli
18	206.5	4.3	1381	3	US-09-540-245A-16	Sequence 16, Appli
19	200	4.2	1069	4	US-09-877-730-2	Sequence 2, Appli
20	200	4.2	1150	4	US-09-877-730-8	Sequence 8, Appli
21	199	4.2	1266	4	US-08-506-296B-4	Sequence 4, Appli
22	196	4.1	904	4	US-09-877-730-6	Sequence 6, Appli
23	196	4.1	985	4	US-09-877-730-10	Sequence 10, Appli
24	193	4.0	380	4	US-09-877-730-4	Sequence 4, Appli
25	186	3.9	1297	3	US-09-540-245A-17	Sequence 17, Appli
26	185	3.9	481	4	US-09-130-491-8	Sequence 8, Appli
27	175.5	3.7	905	4	US-09-369-364A-9	Sequence 9, Appli
28	172	3.6	837	4	US-09-122-126B-2	Sequence 2, Appli
29	172	3.6	837	4	US-09-634-286A-2	Sequence 2, Appli
30	170.5	3.6	1224	4	US-09-930-872-4	Sequence 4, Appli
31	167.5	3.5	757	4	US-09-963-791-24	Sequence 24, Appli
32	167.5	3.5	908	4	US-09-963-791-2	Sequence 2, Appli
33	163	3.4	1081	4	US-09-369-364A-17	Sequence 17, Appli
34	161	3.4	1501	2	US-08-447-464-3	Sequence 3, Appli
35	161	3.4	1501	2	US-08-716-679-3	Sequence 3, Appli
36	160	3.3	930	4	US-09-369-364A-2	Sequence 2, Appli
37	160	3.3	930	4	US-09-122-126B-15	Sequence 15, Appli
38	160	3.3	930	4	US-09-634-286A-15	Sequence 15, Appli
39	159.5	3.3	551	4	US-09-130-491-16	Sequence 16, Appli
40	159.5	3.3	727	4	US-09-445-023A-12	Sequence 12, Appli
41	159.5	3.3	997	4	US-09-369-364A-7	Sequence 7, Appli
42	159	3.3	518	4	US-09-369-364A-22	Sequence 22, Appli
43	158.5	3.3	1911	1	US-08-348-006B-5	Sequence 5, Appli
44	158.5	3.3	1911	2	US-08-800-825A-5	Sequence 5, Appli
45	158.5	3.3	1911	3	US-09-158-657-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1

US-08-808-982-5

; Sequence 5, Application US/08808982

; Patent No. 5939271

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; APPLICANT: Leonardo, E. David

; APPLICANT: Hink, Lindsay

; APPLICANT: Masu, Masayuki

; APPLICANT: Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 898 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-808-982-5

US-08-808-982-5

Qy	421 TIQPDLSTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFVS	480
	: :	
Db	421 TIQPDLSTTTTYQGSLCSRQDGSPKFQLSGHLLSPLGSGRHTLHHSSPTSEAEDFVS	480
Qy	481 RLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMPINTGISLLIPPDAlPRGKIYEIYLTLHK	540
Db	481 RLSTQNYFRSLPRGTSNMAYGTFNFLGGRLMPINTGISLLIPPDAlPRGKIYEIYLTLHK	540
Qy	541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCEGSW	600
	:	
Db	541 PEDVRLPLAGCQTLLSPVVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCEGSW	600
Qy	601 EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLFAPVACT	660
	: : :	
Db	601 EDVLHLGEEPSHLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLLFA PVACT	660
Qy	661 SLEYNIRVYCLHDTHDALKEVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDV PSSLW	720
Db	661 SLEYNIRVYCLHDTHDALKEVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDV PSSLW	720
Qy	721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDDLACKLWVWQVEGDGQSF SINF	780
	: : : : : : :	
Db	721 KSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTSDLACKVWWQVEGDGQSF NINF	780
Qy	781 NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPCCR GADWRTLAQKL	840
	: : : : :	
Db	781 NITKDTRFAELLALESEGGVPALVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQKL	840
Qy	841 HLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC	898
	: : : :	
Db	841 HLD SHLSFFASKPSPTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC	898

RESULT 2

US-09-306-902A-5

; Sequence 5, Application US/09306902A

; Patent No. 6277585

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; Leonardo, E. David

; Hink, Lindsay

; Masu, Masayuki

; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/306,902A
; FILING DATE: 07-May-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 898 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-306-902A-5

Query Match 96.8%; Score 4638; DB 3; Length 898;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 862; Conservative 17; Mismatches 19; Indels 0; Gaps 0;

QY 1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

QY 61 VLLVCKAVPATQIFFKCNGEWRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVFGL 120
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 VLLVCKAVPATQIFFKCNGEWRQVDHVIERSTDSSGLPTMEVRINVSRRQQVEKVFGL 120

QY 121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

QY 181 VEWLRNEDLVPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVIVY 240
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 VEWLRNEDLVPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSTSAAIVIVY 240

QY 241 VNNGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGS 300
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 VNNGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGS 300

QY 301 WSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTLDTRNCTS DLCVHSASGPEDVA 360
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 WSSWSKWSACGLDCTHWRSRECSDPAPRNGGEECGRGADLDTRNCTS DCLCLHTASCPEDVA 360

QY 361 LYVGLIAVAVCLVLLLVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPILL 420
||| :||| :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 LYIGLVAVAVCLFLLLALGLTYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPILL 420

QY 421 TIQPDLSTTTTYQGSCLCPRQDGSPPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFVS 480
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :|||
Db 421 TIQPDLSTTTTYQGSCLCSRQDGSPPKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEDFVS 480

Qy 481 RLSTQNYFRSLPRGTSNMTRYGTFNFLGGRLMPINTGISLLIPPAIPRGKIYEIYLTLHK 540
|||
Db 481 RLSTQNYFRSLPRGTSNMAYGTFNFLGGRLMPINTGISLLIPPAIPRGKIYEIYLTLHK 540

Qy 541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSW 600
|||:
Db 541 PEDVRLPLAGCQTLLSPVVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSW 600

Qy 601 EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLLFAFPVACT 660
|||:
Db 601 EDVLHLGEEPSHLYYCQLEAGACYVFTEQLGRFALVGEALSVAAKRLRLLLFAFPVACT 660

Qy 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW 720
|||
Db 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW 720

Qy 721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDDLACKLWWVQVEGDGQSFINSF 780
|||:
Db 721 KSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTSDLACKVWWVQVEGDGQSFNINF 780

Qy 781 NITKDTRFAELLALESAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKL 840
|||:
Db 781 NITKDTRFAELLALESAGGPALVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQKL 840

Qy 841 HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 898
|||
Db 841 HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC 898

RESULT 3

US-08-808-982-6

; Sequence 6, Application US/08808982

; Patent No. 5939271

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, E. David
; APPLICANT: Hink, Lindsay
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/808,982

; FILING DATE:

; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide

US-08-808-982-6

Query Match 58.8%; Score 2815.5; DB 2; Length 557;
Best Local Similarity 96.8%; Pred. No. 5.2e-259;
Matches 539; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

Qy 343 NCTSDLVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 402
Db 1 NCTSDLXVHTASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 60

Qy 403 GFQPVSIKPSKADNPHLLTIQPDLSSTTTYQGSLCPRQDGSPKFQLTNGHLLSPLGGG 462
Db 61 GFQPVSIKPSKADNPHLLTIQPDLSSTTTYQGSLCPRQDGSPKFQLTNGHLLSPLGGG 120

Qy 463 RHTLHHSSPTSEAEFFSRLSTQNYFRSLPRGTSNMTYGTNFNLGGRLMIPNTGISLLIP 522
Db 121 RHTLHHSSPTSEAEFFSRLSTQNYFRSLPRGTSNMTYGTNFNLGGRLMIPNTGISLLIP 180

Qy 523 PDAIPRGKIYETIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGE 582
Db 181 PDAIPRGKIYETIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGE 240

Qy 583 SPDSWSLRKKQSCEGSWEDVLHLGEEAPSHYYCQLEASACYVFTEQLGRFALVGEALS 642
Db 241 SPDSWSLALKKQSCEGSWEDVLHLGEEAPSHYYCQLEASACYVFTEQLGRFALVGEALS 300

Qy 643 VAAAKRLKLLLFAFPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFK 702
Db 301 VAAAKRLKLLLFAFPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFK 360

Qy 703 DSYHNLRLSIHDVPSLWKSLLVSYQEIPFYHIWNGTQRYLHCTTLERVSPSTSDLAC 762
Db 361 DSYHNLXLSXHDVPSLWKSLLVSYQEIPFYHIWNGTQRYLHCTTLERVSPSTSDLAC 420

Qy 763 KLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISS 822
Db 421 KLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISS 480

Qy 823 LDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAG 882
Db 481 LDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAG 540

Qy 883 LGQPDAGLFT-VSEAEC 898
| : |||||
Db 541 TXPAGRWLLSQCSEAEC 557

RESULT 4
US-09-306-902A-6
; Sequence 6, Application US/09306902A
; Patent No. 6277585
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/306,902A
; FILING DATE: 07-May-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-306-902A-6

Query Match 58.8%; Score 2815.5; DB 3; Length 557;
Best Local Similarity 96.8%; Pred. No. 5.2e-259;
Matches 539; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

Qy 343 NCTSDLCVHSASGPEDVALYVGLIAAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 402
| : ||||| ||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
Db 1 NCTSDLXVHTASGPEDVALYVGLIAAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 60

Qy	403	GFQPVSIKPSKADNPPLLTIQPDLS TTTYQGSCLPRQDGSPKFQLTNGHLLSPLGGG	462
Db	61	GFQPVSIKPSKADNPPLLTIQPDLS TTTYQGSCLPRQDGSPKFQLTNGHLLSPLGGG	120
Qy	463	RHTLHHSSPTSEAEFFVSRLSTQNYFRSLPRTGSNMTYGTNFNLGGRLMIPNTGISLLIP	522
Db	121	RHTLHHSSPTSEAEFFVSRLSTQNYFRSLPRTGSNMTYGTNFNLGGRLMIPNTGISLLIP	180
Qy	523	PDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPEP	582
Db	181	PDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPEP	240
Qy	583	SPDSWSLRKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALS	642
Db	241	SPDSWSLALKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALS	300
Qy	643	VAAAKRLKLLLFAFPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFK	702
Db	301	VAAAKRLKLLLFAFPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHLX	360
Qy	703	DSYHNLRLSIHDPVSSLWKSKLIVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSIDLAC	762
Db	361	DSYHNLXLSXHDVPSSLWKSKLIVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSIDLAC	420
Qy	763	KLWVWQVEGDGQSFISINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISS	822
Db	421	KLWVWQVEGDGQSFISINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISS	480
Qy	823	LDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAG	882
Db	481	LDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAG	540
Qy	883	LGQPDAGLFT-VSEAEC	898
	:		
Db	541	TXPAGRWLSQLQCSEAEC	557

RESULT 5

US-08-808-982-7

; Sequence 7, Application US/08808982

; Patent No. 5939271

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; APPLICANT: Leonardo, E. David

; APPLICANT: Hink, Lindsay

; APPLICANT: Masu, Masayuki

; APPLICANT: Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 943 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide

US-08-808-982-7

Query Match 53.7%; Score 2571.5; DB 2; Length 943;
Best Local Similarity 53.3%; Pred. No. 2.2e-235;
Matches 504; Conservative 142; Mismatches 221; Indels 79; Gaps 16;

Qy 9 PALLGIVLAAWI.RGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLVCKAV 68
|:| || | ||| : : | | : |||||:||| |||||||| | :|
Db 21 PSLAGI-----DSGAQ---GLPDSFPSAPAEQLPHFLLEPEDAYIVKNKPVELHCRAF 70

Qy 69 PATQIFFKCNGEWRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVFGLEEYWCQCVA 128
| |||:| ||||| | || : | | :|| | |||||:| |||:| |||||
Db 71 PATQIYFKCNGEWRQKGHVTQESLDEATGLRIREVQIEVSRRQQVEELFGLEDYWCQCVA 130

Qy 129 WSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCTPPEGIPPAEVWLRNED 188
| |||||:| ||||| | |||:| ||||| | : | :| | |||||:| | |||||:| |||
Db 131 WSSSGTTKSRRAYIRIAYLRKNFDQEPLAKEVPLDHEVLLQCRPPEGVPVAEVWLNED 190

Qy 189 LVDPSLDPNVYITREHSLVVRQARLAQDTANYTCVAKNIVARRRSASAIVIVVNGGWSTW 248
:||: | | :| :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 191 VIDPAQDTNFLLTIDHNLIIQARLSDTANYTCVAKRRTTATVIVVVNGGWSSW 250

Qy 249 TEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWSKWS 308
| || | | |||||:|:| ||||| | ||| | :| |||||:|:| |||||
Db 251 AEWSPCSNRGCRGWQKRTRTCTNPAPLNGGAFCEGQACQKTAJTVCPVDGAWEWSKWS 310

Qy 309 ACGLDCTHWRSRECSDPAPRNGGECCQGTDLDRNCTSDLCV---HSASGP----- 357
| | :| | ||||| | | :|| | :| | |||:| | || | :|:
Db 311 ACSTECAHWRSCRECMAPPQNNGRDCSGTLLDSKNCTDGLCVLNQRTLNDPKSRPLEPSG 370

Qy 358 DVALYVGL-IAVAVCLVLLLVLILVYCRKKEGLDSDVADSS-ILTSGFQPVSIKPSKAD 415
| ||| | | :| | | :|:| | :| | | :|:| | | | | | | :| :|:
Db 371 DVALYAGLVAVFVVLAVLMAVGIVYRRNCRDFDTIDSSAALTGGFPVNFKTARPS 430

QY	416 NPHLL--TIQPDLSTTTTYQGSLCPRQDGSPPKFQLTNGHLLSPL----- 459 : : : : :
Db	431 NPQLLHPSAPPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLLDPLSLKIKVYDSSTIG 489
Qy	460 -GGG-----RHTLHHSSPTSEAEFVSRSLSTQNYFRSLPRGT 495 : : : : :
Db	490 SGAGLADGADLLGVLPVGTYPGDFSRDTHFLHLRS-----ASLGSQ-HLLGLPRDP 539
Qy	496 SNMTYGTFNFLGGRLMIPNTGISLILLPPDAIPRGKIYEIYLTLHKPEDVRLPLA-GCQTL 554 : : : :: :
Db	540 SSSVSGTFGCLGGRLTIPGTGVSSLVPNGAPQGKFYDLYLRINKTEST-LPLSEGSQTV 598
Qy	555 LSPIVSCGPPGVLLTRPVILAMDHCGEPSWLSRLKKQSCEGSWEDVLHLGEEAPSHL 614 : : : : : : : :
Db	599 LSPSVTCGPTGLLCRPVVLTVPHCAEVIAQGDWIFQLKTQAHQGHWEETVLDDEETLNTP 658
Qy	615 YYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLLFAFPVACTSLEYNIRVYCLHDT 674 : :: : :: : : : :: : ::
Db	659 CYCQLEAKSCHILLDQLGTYVFTGESYSRSAVKRLQLAIFAPALCTSLEYSLRVYCLEDT 718
Qy	675 HDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSLWKSLLVSYQEIPFY 734 : :
Db	719 PAALKEVLELERTLGGLVEEPKTLFKDSYHNLRLSLHDIPHAWRSKLLAKYQEIPFY 778
Qy	735 HIWNGTQRYLHCTFTLERVSPSTSSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLAL 794 : : :
Db	779 HVWNGSQKALHCTFTLERHSLASTEFTCKVCVRQVEGEGQIFQLHTTLA-ETPAGSLDAL 837
Qy	795 ESEAGVPAL--VGPSAFKIPFLIRQKIISSLDPCCRQADWRTLAQKLHLDHSFFASK 852 :
Db	838 CSAPGNAATTQLGPYAFKIPPLSIRQKICNSLDAPNSRGNDWRLLAQKLSMDRYLNYFATK 897
Qy	853 PSPTAMILNWEARHFPGNQLSQLAAAVAGLGQPDAGLFTVSEAEC 898 : : : : : : : :
Db	898 ASPTGVILDWEARQQDDGDLNSLASALEEMGKSEMLVAMTTDGDC 943

RESULT 6

US-09-306-902A-7

; Sequence 7, Application US/09306902A

; Patent No. 6277585

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104

; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/306,902A
FILING DATE: 07-May-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC96-217
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 943 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-306-902A-7

```

Query Match          53.7%;  Score 2571.5;  DB 3;  Length 943;
Best Local Similarity  53.3%;  Pred. No. 2.2e-235;
Matches 504;  Conservative 142;  Mismatches 221;  Indels 79;  Gaps 16;

Qy      9 PALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAV 68
       ||:| || | |||| : : | | : ||||:|||| ||||||| | :|
Db     21 PSLAGI-----DSGAQ---GLPDSFPSAPAEQLPHFLLEPEDAYIVKNKPVELHCRAF 70

Qy      69 PATQIFFKCNGEWRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVGLEEYWCQCVA 128
       |||||:|||||| | || : || :|| ||:| |||||||:|||:||||||| |
Db    71 PATQIYFKCNGEWWQSQKGHVTQESLDEATGLRIREVQIEVSRRQQVEELFGLEDYWCQCVA 130

Qy      129 WSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVWLRNED 188
       |||||||:|||||| | |||:|||||:||||||| |: ::| |||||||:| |||||||:|||
Db   131 WSSSGTTKSRRRAYIRIAYLRKNFDQEPLAKEVPLDHEVLLQCRPPEGVPVAEVWLRNED 190

Qy      189 LVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVIVYVNGGWSTW 248
       ::||: | | :| :||:|||||:|||||||:|||||:||| :| |||||||:||| |
Db   191 VIDPAQDTNFLLTIDHNLIIQRQLSDTANYTCVAKNIVAKRSTTATVIVYVNGGWSSW 250

Qy      249 TEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNQVQKTACATLCPVDGWSWPWSKWS 308
       ||| || | |||||:|:|||||||:||||| | ||| | :|||:|:| |||||
Db   251 AEWSPCSNRGGRGWQKRTRTCTNPAPLNGGAFCEGQACQKTACTVCPVDGAWTEWSKWS 310

Qy      309 ACGLDCTHWRSRECSDPAPRNGGEECQGTLDTRNCTS DLCV---HSASGPE----- 357
       || :| | ||||| | | :|| :| || ||:||| ||| : :|:
Db   311 ACSTECAHWSRECMAPPQNNGRDCSGTLLDSKNCTDGLCVLNQRTLNDPKSRPLEPSG 370

Qy      358 DVALYVGL-IAVAVCLVLLLVLILVYCRKKEGLDSDVADSS-ILTSGFQPVSIKPSKAD 415
       ||||| || :|| | | :|: | :|| | | :|: | || | ||| ||| :| :|
Db   371 DVALYAGLVAVFVVLAVLMAVGIVYRRNCRDFDTDIDSSAALTGGFHPVNFKTARPS 430

```

Qy	416 NPHLL--TIQPDLSTTTTYQGSLCPRQDGSPPKFQLTNGHLLSPL----- 459 : : : : :
Db	431 NPQLLHPSAPPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLLDPLPSLKIKVYDSSTIG 489
Qy	460 -GGG-----RHTLHHSSPTSEAEVFVSRNSTQNYFRSLPRGT 495 : : :
Db	490 SGAGLADGADLLGVLPPIPGDFSRDTHFLHLRS-----ASLGSQ-HLLGLPRDP 539
Qy	496 SNMTYGTNFNLGGRLMIPNTGISLLIIPPDAIPRGKIYEIYLTLHKPEDVRLPLA-GCQTL 554 : : : : :: : :
Db	540 SSSVSGTFGCLGGRLTIPGTGVSLVPNGAI PQGKFYDLYLRINKTEST-LPLSEGSQTV 598
Qy	555 LSPIVSCGPPGVLLTRPVILAMDHCGEPSWLSRLKKQSCEGSWEDVLHLGEEAPSHL 614 : : : : : : : : :
Db	599 LSPSVTCGPTGLLLCPVVLTVPHCAEVIAGDWIFQLKTQAHQGHWEVVTLDEETLNTP 658
Qy	615 YYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLFAPVACTSLEYNIRVYCLHDT 674 : : : : : :
Db	659 CYCQLEAKSCHILLDQLGTYVFTGESYSRSAVKRLQLAIFAPALCTSLEYSLRVYCLEDT 718
Qy	675 HDALKEVVQLEKQLGGQLIQEPRVLFHFKDSYHNLRLSIHDPSSLWKSKLVSYQEIPFY 734 : : : : : : :
Db	719 PAALKEVLELERTLGGYLVEEPKTLFKDSYHNLRLSLHDIPHAWRSKLLAKYQEIPFY 778
Qy	735 HIWNGBTQRYLHCTFTLERVSPSTSSDLACKLWWVQVEGDQSFISINFNITKDTRFAELLAL 794 : : : ::
Db	779 HVWNGSQKALHCTFTLERHSLASTEFTCKVCVRQVEGEGQIFQLHTTLA-ETPAGSLDAL 837
Qy	795 ESEAGVPAL--VGPSAFKIPFLIRQKIISSLDPCCRGADWRTLAQKLHLDHSFFASK 852 : :
Db	838 CSAPGNAATTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRLLAQKLSMDRYLNYPATK 897
Qy	853 PSPTAMILNWEARHFPGNQLSQLAAAVAGLGQPDAGLFTVSEAEC 898 : : : : :
Db	898 ASPTGVILDWEARQQDDGDLNSIASALEEMGKSEMLVAMTTDGDC 943

RESULT 7

US-08-313-288B-19

; Sequence 19, Application US/08313288B

; Patent No. 5750502

; GENERAL INFORMATION:

; APPLICANT: Jessell, Thomas M. and Avihu Klar

; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A

; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,288B
FILING DATE: January 5, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526
TELEX:
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1172 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-313-288B-19

RESULT 8
US-08-808-982-8
; Sequence 8, Application US/08808982
; Patent No. 5939271
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, E. David
; APPLICANT: Hink, Lindsay
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide

US-08-808-982-8

Query Match 6.1%; Score 294; DB 2; Length 102;
Best Local Similarity 56.4%; Pred. No. 3.9e-20;
Matches 57; Conservative 16; Mismatches 28; Indels 0; Gaps 0;

Qy 608 EEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSAAKRLKLLFAPVACTSLEYNIR 667
|| : | ||| || : :||| : ||: | : |||:| :||| |||||:|||
Db 2 EETLNTPCYXQLEPRACXILLDQLGTVFTGESYSRSAVKRLQLAVFAPALCTSLEYSLR 61

Qy 668 VYCLHDTDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNL 708
|||| || |||||:|||: ||| |::||: | |||||||
Db 62 VYCCLEDTPVALKEVLERTLGGYLVEEPKPLMFKDSYHNL 102

RESULT 9
US-09-306-902A-8
; Sequence 8, Application US/09306902A
; Patent No. 6277585
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/306,902A
; FILING DATE: 07-May-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-306-902A-8

Query Match 6.1%; Score 294; DB 3; Length 102;
Best Local Similarity 56.4%; Pred. No. 3.9e-20;
Matches 57; Conservative 16; Mismatches 28; Indels 0; Gaps 0;

Qy	608	EEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLL	FAPVACTSLEYNIR	667	
	:	: : : : : : :	::		
Db	2	EETLNTPCYXQLEPRACXILL	DQLGTYVFTGESYSRS	AVKRLQLAVFAPALCTSLEYSLR	61

Qy	668	VYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNL	708		
		:: : :: :			
Db	62	VYCLEDTPVALKEVLELERTLG	GGYLVEEPKPLMF	KDSYHNL	102

RESULT 10
PCT-US93-01652-1
; Sequence 1, Application PC/TUS9301652
; GENERAL INFORMATION:
; APPLICANT: Bouck, Noel P.
; APPLICANT: Polverini, Peter J.
; APPLICANT: Good, Deborah J.
; APPLICANT: Frazier, William A.
; TITLE OF INVENTION: Method and Composition for

; TITLE OF INVENTION: Inhibiting Angiogenesis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Tilton, Fallon, Lungmus & Chestnut
; STREET: 100 South Wacker Drive, Suite 960
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/01652
; FILING DATE: 19930222
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/841,656
; FILING DATE: 24-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/464,369
; FILING DATE: 12-JAN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Fentress, Susan B.
; REGISTRATION NUMBER: 31,327
; REFERENCE/DOCKET NUMBER: 92005-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)-456-8000
; TELEFAX: (312)-456-7776
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 239 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide

PCT-US93-01652-1

Query Match 5.6%; Score 268.5; DB 5; Length 239;
Best Local Similarity 33.5%; Pred. No. 4.4e-17;
Matches 52; Conservative 23; Mismatches 61; Indels 19; Gaps 4;

QY 207 VVRQARLADTANYTCVAKNIVAR-----RRSASAIVIVYVNGGWSTWTEWSVCSASC 258
Db : :: | | : | | :: | :| :| || | :|| | :| || | :|
QY 259 GRGWQKRSRSCTNPAPLNGGAFCEGQNQKTAC-ATLCPVDGSWPWSKWSACGLDC--- 314
Db | | | | | :| :| | | || :| :| || | :| | | | :|
QY 141 GDGVITRIRLCNSPSPQMNGLPCEGEARETKACKDACPINGGWGPSPWDICSVTCGGG 200
Db
QY 315 THWRSRECSDPAPRNGGEECQGTDLDTRNCTS DLC 349
Db ||| |::|||: | | :| | :| :| |
Db 201 VQKRSRLCNNPAPQFGLDCVGDTVENQICNKQDC 235

RESULT 11
US-08-313-288B-20
; Sequence 20, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

US-08-313-288B-20

Query Match 5.6%; Score 268.5; DB 1; Length 1170;
Best Local Similarity 32.9%; Pred. No. 6.3e-16;
Matches 51; Conservative 24; Mismatches 61; Indels 19; Gaps 4;

QY	207 VVRQARLADTANYTCVAKNIVAR-----RRSASAIVYVNNGGSTWTEWSVCASC 258
	: :: : :: : : : :
Db	399 IQQQRGRSCDSLNNRCEGSSVQTRTCHIQECDKRFKQ-----DGGWSHWSPWSSCSVTC 451
QY	259 GRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTAC-ATLCPVDGSPWSKWSACGLDC-- 314
	: : : : :
Db	452 GDGVITRIRLNCNSPSPQMNGKPCEGEARETKACKDACPINGGWGPWSPWDICSVTCGGG 511
QY	315 THWRSRECSDPAPRNGGEECQGTLDTRNCTSDL 349
	:: : :: : :
Db	512 VQKRSRLCNNPTPQFGGKDCVGDVTEENQICNKQDC 546

RESULT 12
US-08-985-526-3
; Sequence 3, Application US/08985526
; Patent No. 6080728
; GENERAL INFORMATION:
; APPLICANT: Mixson, James A
; TITLE OF INVENTION: CARRIER:DNA COMPLEXES CONTAINING DNA
; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN
GENE
; TITLE OF INVENTION: THERAPY
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Connolly, Bove, Lodge, & Hutz
; STREET: 1220 Market Street, P.O. Box 2207
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: U.S.A.
; ZIP: 19899
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,526
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,845
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McMorrow Jr., Robert G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 658-9141
; TELEFAX: (302) 658-5613
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-985-526-3

Query Match 5.2%; Score 249.5; DB 3; Length 441;
 Best Local Similarity 26.2%; Pred. No. 7.9e-15;
 Matches 88; Conservative 35; Mismatches 112; Indels 101; Gaps 16;

Qy 114 EKVGLEEYWCQCVAWSGGTTSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPP 173
 : | | | | | : | | | : || | : | | |
 Pb 182 NGCWGPRWSRWDIGCSVTGGGGVQKDSPLCV DSPMTTEENKEIANELP PPR 229

Ov 174 ----EGTTPAEEVWLBNED-LYDPSLDPNVYTREHSLVYVBOARLADTANYTCVAKNLYA 228

| | :| :|| :| :| :| :|
Db 240 LCYHNG----VQYRNNEEWTVDSCTE-----CHCQNSVT 269

Qy 229 RRRSASAIVIVYVNG-----GWSTWTEWSVCASCGRGWQKRSRSC 269
: | :| :| :| :| :| :| :| :| :|
Db 270 ICKKVSCPIMPCSNAVPDGECCPRCWPSDSADDGWPWSEWTSCSTSCGNGIQQRGRSC 329

Qy 270 TNPAVLNGGAFCEGQNVQKTAC-ATLC---PVDGSWSPWWSKWSACGLDC---THWRSRE 321
: ||| :|| | | | | ||| ||| :| :| :|
Db 330 DS---LNRR--CEGSSVQTRTCHIQECDKRFKQDGWWSHWPSSCSVTCGDGVITRITL 384

Qy 322 CSDPAPRNGGEECQGTLDTRNCTSDL-C-VHSASGP 356
| :| :| :| :| :| :| :| :| :|
Db 385 CNSPSPQMNGKPCEGEARETKACKDACPINGGWGP 420

RESULT 13

US-08-313-288B-15

; Sequence 15, Application US/08313288B

; Patent No. 5750502

; GENERAL INFORMATION:

; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US

TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

; LENGTH: 469 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO


```

Query Match           5.0%; Score 238; DB 3; Length 218;
Best Local Similarity 39.3%; Pred. No. 3e-14;
Matches 48; Conservative 16; Mismatches 44; Indels 14; Gaps 6;

Qy      244 GWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNNGAFCEGQNVQKTAC-ATLC---PVD 298
        ||| |:||: || ||| | |:| ||| : ||| ||| :|| | | | |
Db      81  GWSPWSEWTSCSTSCGNGIQQRGRSCDS---LNNR--CEGSSVQTRTCHIQECDKRFKQD 135

Qy      299 GSWSPWSKWSACGLDC---THWRSRECSDPAPRNGGECCQGTDLDTRNCTS DLC-VHSAS 354
        | ||| |||:| :| | | | |:|: |:| :|:|:| | | | :|
Db      136 GGWSHWSPWSSCSVTCGDGVITRITNLCS PSQMNGKPCEGRE AETKACKKDACPINGW 195

Qy      355 GP 356
        ||
Db      196 GP 197

```

RESULT 15

US-09-540-245A-15

; Sequence 15, Application US/09540245A

; Patent No. 6270984

; GENERAL INFORMATION:

; APPLICANT: Goodman, Corey

; APPLICANT: Kid, Thomas

; APPLICANT: Brose, Katja

; APPLICANT: Tessier-Lavigne
TITLE OF INVENTION: Model

; TITLE OF INVENTION: Modulating Robo: Ligand
FILE REFERENCE: P00-021-2

; FILE REFERENCE: B98-031-3
; CURRENT APPLICATION NUMBER: US/0

CURRENT APPLICATION NUMBER: US/09/5
CURRENT FILING DATE: 2000-03-31

PBTOR APPLICATION NUMBER: 60/0

PRIOR APPLICATION NUMBER: 08/685,511

PRIOR APPLICATION NUMBER: 60/0

PRIOR FILING DATE: 1998-04-07

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn Ver. 2

; SEQ ID NO 15

; LENGTH: 13

; TYPE: PRT

; ORGANISM:

US-09-540-245A-15

Best Local Similarity 20.7%; Pred. No. 1.5e-12;

Matches 187; Conservative 104; Mismatches 273

	PROTEIN	SEQUENCE	PROTEIN	SEQUENCE
Db	28	RMWLLPAWLLLVLVA-----	SNGLPAVRGQYQSPRIIEH---	PTDLVVKKN 70
Qy	59	KPVLLVCK--AVPATQIFFKCNGEWV---RQVDHVIERSTDGSSGLPTMEVRINVSRQQV	11	
Db	71	EPATLNCKVEGKPEPTIEWFKDGEPVSTNEKKSHRVQFKDGALFFYRTM-----	QG 12	
Qy	114	EKFVFGLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPP	17	

Db	122 KKEQDGGEYW--CVAKNRVGQAVSRHASLQIAVLRDDFRVEPKDTRVAKGETALLECGPP	179
Qy	174 EGIPPAEVIEWLRN---EDLVDSL--DPNVYITREHSLVVRQARLADTANYTCVAKNIV	227
	: : ::: : : ::: : : :	
Db	180 KGIPPEPTLIWIKDGVPPLDLKAMSGASSRVRIVDGGNLLISNVEPIDEGNYKCIAQNLV	239
Qy	228 ARRRSASAAVIVYVN-----GGWSTWTEWSVCSASCGRG-----WQK-	264
	: : : : :	
Db	240 GTRESSYAKLIVQVKPYFMKEPKDQVMLYGQTATF-----HCSVGGDPPPVLWKKE	291
Qy	265 -----RSRSCTNPAPLNNGAF-CEGQN--VQKTACATL-----	294
	: : : : :	
Db	292 EGNIPVSRARILHDEKSLEISNITPTDEGTYVCEAHNNVGQISARASLIVHAPPNFTKRP	351
Qy	295 -----CPVDGSWSPWSKWSACGL-----DCTHWRSRECSDPAPRNGGEEC	334
	: : : : :	
Db	352 SNKKVGLNGVVQLPCMASNPPPSVFWTKEGVSTLMFPNSSHGRQYVAADGT-----L	404
Qy	335 QGTDLDTRNCTS DL CVHSASGPEDVALYVGLIAVAVCLVLLLVLILVYCRKKEGLDSV	394
	: :	
Db	405 QITDV-----RQEDEGYY-----VCSAF-----SV	424
Qy	395 ADSSILTSGFQPVSIKPSKADN--PHLLTIQPDLS TTTTYQGSL---CPRQDGSPSKF	448
	: : : :	
Db	425 VDSSTVR-----VFLQVSSVDERPPPIIQIGP---ANQTLPKGSVATLPCRATGNPSPRI	476
Qy	449 Q-LTNGHLLSPLGGGRHTLHHSSPTSEAEFFVSRLSTQNYFRSLPRGTSNMTYGTNFLG	507
	: : : : : :	
Db	477 KWFHDGHAVQ--AGNRYSIIQGSSLRVDDLQLSDSGTYTCTASGERGETS-----	524
Qy	508 GRLMIPNTGISLLIPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLSPIVSCGPPGVL	567
	: :	
Db	525 -----WAATLTVEKPGSTSLHRAA-----DPSTYPAPP GT-	554
Qy	568 LTRPVILAMDHCGEPSDWSLR-LKKQSCEGS-----WEDVLH-L	606
	: :	
Db	555 ---PKV LNV-----SRTSISLRWAKSQEKGAVGPIIGYTVEYFSPDLQ TGWIVAAHRV	605
Qy	607 GEEAPSHLYYCQLEASACYVF-----TEQLGRFALVGEALSVA-	644
	: :	
Db	606 GD---TQVTISGLTPGTSYFLVRAENTQGISVPSGLSNVIKTIEADF DAASANDLSAAR	662
Qy	645 ---AAKRLKLLL FAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLI QEPRVLHF	701
	: : : : : : : : :	
Db	663 TLLTGKSVELIDASAINASAVRLE---WMLHVSAD-----EKYVEGLRI-----HY	705
Qy	702 KDS-----YHNRL-----SIHDVPSLWKS KLLVS	727
	:	
Db	706 KDASVPSAQYHSITVMASAESFV VGNLK KYTKYEFFLTPFFETIEGQPSN---SKTALT	762
Qy	728 YQEIP 732	
	:::	
Db	763 YEDVP 767	

Search completed: July 6, 2004, 14:36:12
 Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2004, 14:30:11 ; Search time 27 Seconds
(without alignments)
3199.259 Million cell updates/sec

Title: US-10-624-932-2

Perfect score: 4791

Sequence: 1 MAVRPGLWPALLGIVLAAWL.....AVAGLGQPDAGLFTVSEAEC 898

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			Description
		Match	Length	DB	
1	977	20.4	919	2	T32541 unc-5 protein - Ca
2	977	20.4	947	1	B44294 unc-5 protein, lon
3	298.5	6.2	1584	2	T00026 brain-specific ang
4	296.5	6.2	1172	1	TSHUP2 thrombospondin 2 p
5	293	6.1	1074	2	JC5928 semaphorin F precu
6	293	6.1	1172	2	A42587 thrombospondin 2 p
7	276	5.8	1444	2	T18856 angiogenesis inhib
8	275	5.7	984	2	T00326 hypothetical prote
9	275	5.7	1522	2	T00028 brain-specific ang
10	274.5	5.7	1572	2	T00027 brain-specific ang
11	270.5	5.6	1170	2	A40558 thrombospondin 1 p
12	268.5	5.6	1170	1	TSHUP1 thrombospondin 1 p
13	263	5.5	1178	1	A39804 thrombospondin pre

14	243	5.1	469	1	S29126	properdin precursor
15	229	4.8	437	2	S05478	properdin - mouse
16	226	4.7	254	2	T15952	hypothetical protein
17	221.5	4.6	1265	1	A37967	neural cell adhesion molecule
18	215	4.5	788	2	T25061	hypothetical protein
19	215	4.5	1651	2	T14160	transmembrane receptor
20	208	4.3	1612	2	T30805	duttl protein - mouse
21	191.5	4.0	1344	2	T14316	rig-1 protein - mouse
22	191	4.0	1863	2	S46217	protein-tyrosine-phosphatase
23	188.5	3.9	957	2	T15976	hypothetical protein
24	186	3.9	423	2	T29549	hypothetical protein
25	181.5	3.8	1273	2	T42405	sax-3 protein - Caenorhabditis elegans
26	181	3.8	1736	2	A47747	tight junction protein
27	178	3.7	1745	2	A46431	tight junction-associated protein
28	175	3.7	1907	2	S50893	protein-tyrosine-phosphatase
29	172	3.6	837	2	T00355	hypothetical protein
30	169.5	3.5	934	1	A34372	complement C6 precursor
31	168.5	3.5	152	2	D89753	protein F11C7.2 [incomplete]
32	168	3.5	860	2	T16892	hypothetical protein
33	162.5	3.4	654	2	T29247	hypothetical protein
34	161	3.4	1501	2	I58148	protein-tyrosine-phosphatase
35	159.5	3.3	951	2	T00017	gene ADAMTS-1 protein
36	159	3.3	805	2	T34212	hypothetical protein
37	158.5	3.3	2165	2	T21371	hypothetical protein
38	157	3.3	550	2	T47158	hypothetical protein
39	156.5	3.3	807	2	A38152	F-spondin - rat
40	156	3.3	584	1	C8HUA	complement C8 alpha
41	155	3.2	1499	2	I50212	protein-tyrosine-phosphatase
42	155	3.2	1898	2	S46216	leukocyte antigen-protein
43	150.5	3.1	1437	2	T31093	probable protein-t
44	149.5	3.1	712	2	A45638	immunodominant microtubule-associated protein
45	148.5	3.1	206	2	A45517	coccidioides-relate

ALIGNMENTS

RESULT 1

T32541

unc-5 protein - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 31-Jan-2000

C;Accession: T32541

R;Latreille, P.

submitted to the EMBL Data Library, December 1997

A;Description: The sequence of *C. elegans* cosmid B0273.

A;Reference number: Z21187

A;Accession: T32541

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-919 <LAT>

A;Cross-references: EMBL:AF036698; PIDN:AAB88355.1; GSPDB:GN00022; CESP:B0273.4a

A;Experimental source: strain Bristol N2; clone B0273

C;Genetics:

A;Gene: unc-5; CESP:B0273.4a

A;Map position: 4

A;Introns: 41/3; 108/1; 142/3; 201/1; 323/2; 553/1; 858/3

C;Superfamily: unc-5 protein; immunoglobulin homology; SH3 homology; thrombospondin type 1 repeat homology

Db	713 LCESNDFILNEKG---NLCICIEDVIPGFSCDGPEVVEISETQHRFV---AQNGLHCSLK	766
Qy	750 LERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAF	809
	: : :: : : : :: : : :	
Db	767 FRPKEINGSQFSTRVIVYQKASSTEPMVM--EVSNPELYDATSEEREKGGSVCV----EF	820
Qy	810 KIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDSHLSFFASKP--SPTAMILNLWEARH	867
	:: :: :: : : : :: :	
Db	821 RLPFGVKDELARLLDMPNESHSDWRGLAKKLHYDRYLQFFASFPDCSPTSLLDLWEASS	880
Qy	868 FPNGN-LSQLAAAVAGLGQPDA	888
	: : : :	
Db	881 SGSARAVPDLLQTLRVMGRPDA	902

RESULT 2

B44294

unc-5 protein, long form - *Caenorhabditis elegans*

N;Contains: unc-5 protein, short form

C;Species: *Caenorhabditis elegans*

C;Date: 30-Apr-1993 #sequence_revision 28-Jul-1995 #text_change 05-Nov-1999

C;Accession: B44294; T32540; A44294

R;Leung-Hagesteijn, C.; Spence, A.M.; Stern, B.D.; Zhou, Y.; Su, M.W.; Hedgecock, E.M.; Culotti, J.G.

Cell 71, 289-299, 1992

A;Title: UNC-5, a transmembrane protein with immunoglobulin and thrombospondin type 1 domains, guides cell and pioneer axon migrations in *C. elegans*.

A;Reference number: A44294; MUID:93046629; PMID:1384987

A;Contents: variety Bergerac

A;Accession: B44294

A;Molecule type: DNA

A;Residues: 1-947 <LEU>

A;Cross-references: GB:S47168; NID:g258527; PIDN:AAB23867.1; PID:g258529

A;Note: sequence extracted from NCBI backbone (NCBIN:116668, NCBIN:116670, NCBIN:116672, NCBIN:116674, NCBIN:116676, NCBIN:116678, NCBIN:116680, NCBIN:116682, NCBIN:116685, NCBIP:118648)

A;Note: authors translated the codon CTA for residue 642 as Val; sequence shown follows the authors' translation

A;Note: mRNA lacking the first exon is equally prevalent

R;Latreille, P.

submitted to the EMBL Data Library, December 1997

A;Description: The sequence of *C. elegans* cosmid B0273.

A;Reference number: Z21187

A;Accession: T32540

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-947 <LAT>

A;Cross-references: EMBL:AF036698; PIDN:AAB88356.1; GSPDB:GN00022; CESP:B0273.4b

A;Experimental source: strain Bristol N2; clone B0273

C;Genetics:

A;Gene: unc-5

A;Map position: 4

A;Introns: 28/1; 69/3; 136/1; 170/3; 229/1; 351/2; 581/1; 886/3

C;Function:

A;Description: required for guidance of pioneering axons and cells migrating dorsally along the body wall; proposed to be a receptor on the surface of the motile cells

C;Superfamily: unc-5 protein; immunoglobulin homology; SH3 homology;
thrombospondin type 1 repeat homology
C;Keywords: alternative splicing; duplication; glycoprotein; receptor;
transmembrane protein
F;30-947/Product: unc-5 protein, short form #status predicted <ALT>
F;46-116/Domain: immunoglobulin homology <IM1>
F;153-211/Domain: immunoglobulin homology <IM2>
F;229-300/Domain: thrombospondin type 1 repeat homology #status atypical <THR1>
F;301-354/Domain: thrombospondin type 1 repeat homology <THR2>
F;365-390/Domain: transmembrane #status predicted <TMM>
F;512-559/Domain: SH3 homology <SH3>
F;53-114, 65-112, 160-209/Disulfide bonds: #status predicted
F;206/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 20.4%; Score 977; DB 1; Length 947;
 Best Local Similarity 28.7%; Pred. No. 1.2e-62;
 Matches 265; Conservative 168; Mismatches 379; Indels 110; Gaps 31;

 49 EPEDVYIVKVNKPVLVCKAVPATQIFFKCNGEWRQVDHVIER--STDGSSGLPTMEVRI 106
 ::|: |:::||: | :| ||:| :||: | :| | ||: | ||:| :||:| :| :|:
 37 QPKSGYVIRNKPLRLQCRAHATKIRYKCSSKWID--DSRIEKLIGTDSTSGVGYIDASV 94

 107 NVSRQQVEKVFGLEEYWCQCVAWSSG-----TTKSQKAYIRIARLRKNFEQEPLAKEVS 161
 ::|| |: :: : ||| | || | | :| :||:| :||:| :|:
 95 DISRIDVDTSGHVDAFQCQCYA--SGDDDQDVVASDVATVHLAYMRKHFLKSPVQAQRVQ 151

 162 LEQGIVLPCRPPEGIPPAEVWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTC 221
 : |||: | | | ||: | :| | | | | | :| :|: | | |:
 152 EGTTLQLPCQAPESDPKAELTWYKDGVVVQP--DANVIRASDGSLIMSAARLSDSGNYTC 209

 222 VAKNIVARRRSASAIVIVYVNGGWSTWTEW-SVCSASC-----RGWQKR 265
 | |: | :: | :||:| | | | | | | | | | | | |:
 210 EATNVANSRKTDPVEQIYDGGWSEWSPWIGTCHVDCPLLQHAHRIRDPHDVLPHQRR 269

 266 SRSCTNPAPLNGGAFCEGQNQVQKTACATLCPVDGSWSPWWSKWSACGLDCTHWRSRECSDP 325
 :||: | | | | | | :||: | :| | | | | | | | | | |:
 270 TRTCNNPAPLNDGEYCKGEEEMTRSCKVPCQLDGGWSSDWSACSSSCHRYSRTRACTVP 329

 326 APRNGGEECQGTDDLTRNCTS DL CVHSASG--PEDVALYVGLIAVAVCLVLLLVLILVY 383
 | | |: | | | | | :| :| | | | | | | | | | | | |:
 330 PPMNGGQPCFGDDLMTQECPAQQLCTADSSRIVISDTAVYGSVASIFIVASFILAILAMFC 389

 384 CR-----KKEGLDSDVADSSILTSGFQPVSIKPSKADNPHLIT----- 422
 |: | :| :| | :| :| :| :| :| :| :| :| :| :|:
 390 CKRGNSKKSKPLKPQKMNSEKAGGIYYS--EPPGVRRLLLEHQHGTLLGEKISSCSQYF 446

 423 -QPDLSTTTT-----TYQGSLCPRDGPSPKFQLTNGHLLSPLGGRHTLHHSSPT-SE 474
 | | :| | :| | | | | | | | | | | | | | | | |:
 447 EPPPLPHSTTLRGKSAFSGSSTRNAGSRAALIQECSSSSGSGGGKRTMLRTSSNCSD 506

 475 AEEFVSRSLSTQNYFRSLPRGTS-NMTYGTNFLLGGRLMPI PNTGISLLIPDAIPRGKIYE 533
 : : : | | | | : | | | | | | | | | | | | | |:
 507 DDNYATLYDYMEDKSVLGLDTSQNIVAAQIDSNGARLSLSKSGARLIVPELAVEGEKM-- 564

 534 IYLTLLHKPEDVRLPLAGCQTLLSPIVSCGPPGV-----LLTRPVILAMDHCGEPS-D 585
 :| | : | :| | | | | | | | | | | | | | | | |:
 565 LYLAVIDTLDQPHLKPIESALSPVIVIGQCDVMSAHDNLRRPVVVSFRHCASTFPRD 624

Qy	586 SWSLRLKKQSCEGS-WEDVLHLGEEAPSHLYYCQLEASA-----CYVFTEQLGRFAL	636
	: : : : : : :	
Db	625 NWQFTL--YADEGSGWQKAVTIGEENLNTNMFVQFEQPGKKNDGFGWCHVMTYSLARML	682
Qy	637 VGEAL--SVAAAKRLKLLLFAFPVACTSLE--YNIRVYCLHDTHDALKEVVQLEKQLGGQL	692
	: : :: : : :	
Db	683 AGHPRRNLSAAKRVHLAVFGPTEMSAYRRPFELRVYCVPETGAAMESVWKQED--GSRL	740
Qy	693 IQEPR--VLHFKDSYHNLRLSIHDV-PSSLWKSKLVLVSYQEIPFYHIWNGTQRYLHCTFT	749
	: : : : :	
Db	741 LCESNDIFILNEKG---NLCICIEDVI PGFSCDGPEVVEISETQHRFV---AQNGLHCSLK	794
Qy	750 LERVSPSTSDLACKLWWVQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAF	809
	: : : : : : : : :	
Db	795 FRPKEINGSQFSTRVIVYQKASSTPMVM--EVSNPELYDATSEEREKGGSVCV---EF	848
Qy	810 KIPFLIRQKISSLDPCCRGADWRTLAQKLHLDHSFFASKP--SPTAMILNLWEARH	867
	: : : : : : : :	
Db	849 RLPFGVKDELARLLDMPNESHSDWRGLAKKLHYDRYLQFFASFPCSPSTSLLDLWEASS	908
Qy	868 FPNGN-LSQLAAVAGLQPD A 888	
	: : : :	
Db	909 SGSARAVPDLLQTLRVMGRPDA 930	

RESULT 3

T00026

brain-specific angiogenesis inhibitor 1 - human

N;Alternate names: BAI1 protein

C;Species: Homo sapiens (man)

C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 12-Feb-1999

C;Accession: T00026

R;Nishimori, H.; Shiratsuchi, T.; Urano, T.; Kimura, Y.; Kiyono, K.; Tatsumi, K.; Yoshida, S.; Ono, M.; Kuwano, M.; Nakamura, Y.

submitted to the EMBL Data Library, June 1997

A;Reference number: Z14064

A;Accession: T00026

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-1584 <NIS>

A;Cross-references: EMBL:AB005297; NID:d1175078; PID:d1024528

A;Experimental source: brain

C;Genetics:

A;Gene: GDB:BAI1

A;Cross-references: GDB:9838088; OMIM:602682

A;Map position: 8q24-8q24

C;Superfamily: thrombospondin type 1 repeat homology

F;408-462/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 6.2%; Score 298.5; DB 2; Length 1584;
 Best Local Similarity 33.5%; Pred. No. 4.4e-13;
 Matches 78; Conservative 35; Mismatches 91; Indels 29; Gaps 11;

Qy	124 CQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEW	183
	: : : : : :	
Db	309 CNREACGPAGRTSSRSQSLRSTDARR--REELGDEL--QQFGFPA-PQTGDPAAE-EW	360

Qy	184 LRNEDLVDPSDLPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVIVNG	243
	: : :: : : : : :	
Db	361 --SPWSVCSSCTCGEGWQTR-----TRFCVSSSYSTQCSGPLREQRLCNNSAVCPVHG	410
Qy	244 GWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNNGAFCEGQNVQKTAC-ATLCP---VDG	299
	: :: :: : :	
Db	411 AWDEWSPWSLCSSCTCGRGFRDRTTCR--PPQFGGNPCGPEKQTFCNIALCPGRAVDG	468
Qy	300 SWSPWSKWSACGLDCT---HWRSRECSDPAPRNGGEECQGTLDTRNCTS D L C	349
	:: : : : : : :: :	
Db	469 NWNEWSSWSACSASC SQGRQORTRECNGPS--YGGAECQGHVETRDCF L QQC	519

RESULT 4

TSHUP2

thrombospondin 2 precursor - human

C;Species: Homo sapiens (man)

C;Date: 19-May-1995 #sequence_revision 03-Aug-1995 #text_change 13-Aug-1999

C;Accession: A47379; A42173

R;LaBell, T.L.; Byers, P.H.

Genomics 17, 225-229, 1993

A;Title: Sequence and characterization of the complete human thrombospondin 2 cDNA: potential regulatory role for the 3' untranslated region.

A;Reference number: A47379; MUID:94010892; PMID:8406456

A;Accession: A47379

A;Molecule type: mRNA

A;Residues: 1-1172 <LAB>

A;Cross-references: GB:L12350; NID:g307505; PIDN:AAA03703.1; PID:g307506

R;LaBell, T.L.; Milewicz, D.J.; Disteche, C.M.; Byers, P.H.

Genomics 12, 421-429, 1992

A;Title: Thrombospondin II: partial cDNA sequence, chromosome location, and expression of a second member of the thrombospondin gene family in humans.

A;Reference number: A42173; MUID:92217961; PMID:1559694

A;Accession: A42173

A;Molecule type: mRNA

A;Residues: 560-1172 <LA2>

A;Cross-references: GB:M81339

A;Experimental source: fibroblast

A;Note: sequence extracted from NCBI backbone (NCBIN:95091, NCBIP:95096)

C;Genetics:

A;Gene: GDB:THBS2; TSP2

A;Cross-references: GDB:128789; OMIM:188061

A;Map position: 6q27-6q27

C;Complex: homotrimer, disulfide linked

C;Function:

A;Description: participates in cell migration and adhesion, and in platelet aggregation

C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology

C;Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer

F;1-18/Domain: signal sequence #status predicted <SIG>

F;19-1172/Domain: thrombospondin 2 #status predicted <MAT>

F;319-377/Domain: von Willebrand factor type C repeat homology <VWC>

F;380-431/Domain: thrombospondin type 1 repeat homology <THR1>

F;436-492/Domain: thrombospondin type 1 repeat homology <THR2>

RESULT 5

JC5928

semaphorin F precursor - human

C;Species: Homo sapiens (man)

C;Date: 10-Apr-1998 #sequence_revision 08-May-1998 #text_change 17-Nov-2000
C;Accession: JC5928

R; Simmons, A.D.; Pu

Biochem. Biophys. Res. Commun. 242, 685-691, 1998
A:Title: Molecular cloning and mapping of human semaphorin F from the Cri-du-

A: Reference number: JG5928; MVID: 98125554; PMID: 9464278

A:REFERENCE NUMBER: UC5928, MOLB:98123334, PMID:9464278
A:Accession #: TC5928

A;Accession: JC5928
A;Status: nucleic a

A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA

A;Molecule type: mRNA
A;Residues: 1-1074 18

A;Cross-references: CR:II

A;Cross-references: GB:052840; NID:g2772583; PID: AAC09473.1; PID:g2772584

A; Experimental source: brain
G:Comments: This section is di-

C;Comment: This protein disrupts normal brain development and leads to some of the features of Cri-du-chat.

C; Genetics:

A;Gene: semaf

C;Superfamily: human semaphorin F; thrombospondin type 1 repeat homology

F;1-20/Domain: signal sequence #status predicted <SIG>
 F;50-533/Domain: semaphorin #status predicted <SEM>
 F;840-896/Domain: thrombospondin type 1 repeat homology <THR3>
 F;971-993/Domain: transmembrane #status predicted <TMM>

Query Match 6.1%; Score 293; DB 2; Length 1074;
 Best Local Similarity 45.8%; Pred. No. 6.6e-13;
 Matches 54; Conservative 11; Mismatches 49; Indels 4; Gaps 2;

Qy	241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGONVQKTACATL-CPVDG 299
	: ::
Db	783 VNGAWSAWTSWSQCSRDCSRGIRNRKRVCNNPEPKYGGMPCLGPSLEYQECNTLPCPVDG 842
Qy	300 SWSPWWSKWSACGLDC---THWRSRECSDPAPRNGGEECQGTLDTRNCTSILCVHSAS 354
	: : : : : :
Db	843 VWSCWSPWTKCSATCGGGHYMRTSNSNPAPAYGGDICLGLHTEALCNTQPCPESWS 900

RESULT 6

A42587

thrombospondin 2 precursor - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
 C;Accession: A42587; A39851
 R;Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.
 J. Biol. Chem. 267, 3274-3281, 1992
 A;Title: Characterization of mouse thrombospondin 2 sequence and expression
 during cell growth and development.
 A;Reference number: A42587; MUID:92147683; PMID:1371115
 A;Accession: A42587
 A;Status: preliminary; not compared with conceptual translation
 A;Molecule type: nucleic acid
 A;Residues: 1-1172 <LAH>
 A;Cross-references: GB:L07803; GB:M87275; NID:g340421; PIDN:AAA53064.1;
 PID:g567241
 A;Note: sequence extracted from NCBI backbone (NCBIP:81502)
 R;Bornstein, P.; O'Rourke, K.; Wikstrom, K.; Wolf, F.W.; Katz, R.; Li, P.;
 Dixit, V.M.
 J. Biol. Chem. 266, 12821-12824, 1991
 A;Title: A second, expressed thrombospondin gene (Thbs2) exists in the mouse
 genome.
 A;Reference number: A39851; MUID:91302287; PMID:1712771
 A;Accession: A39851
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-873 <BOR>
 A;Cross-references: GB:M64866; NID:g201994; PIDN:AAA40432.1; PID:g201995
 C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat
 homology; von Willebrand factor type C repeat homology
 C;Keywords: calcium binding; glycoprotein
 F;319-377/Domain: von Willebrand factor type C repeat homology <VWC>
 F;380-431/Domain: thrombospondin type 1 repeat homology <THR1>
 F;436-492/Domain: thrombospondin type 1 repeat homology <THR2>
 F;493-549/Domain: thrombospondin type 1 repeat homology <THR3>
 F;553-588/Domain: EGF homology <EGF1>
 F;652-691/Domain: EGF homology <EGF>

Query Match 6.1%; Score 293; DB 2; Length 1172;
 Best Local Similarity 38.0%; Pred. No. 7.4e-13;
 Matches 60; Conservative 22; Mismatches 66; Indels 10; Gaps 5;

Qy	209 RQARLADTANYTCVAKNIVARRRS-ASAAVIVYVNGGWSTWTEWSVCASACGRGWQKRSR 267
	:: : : : : : :
Db	403 QRGRSCDVTNSNTCLG PSI QT RTCS LGKCDTRI RQNGGWSHW SPWSSCS VTCGVGNVTRIR 462
Qy	268 SCTNPAPLNGGAFCEGQNVQKTAC-ATLC PVDGSWSPWSKWSACGLDCT---HWRSRECS 323
	: : : : : :
Db	463 LCNSPVPQMGGKNCKSGRET KPCQRDP CPIDGRWSPWSPWSACTVTCAGGIRERSRV CN 522
Qy	324 DPAPRNGGEECQG--TD--LDTRNCTS DLCVHSASGP 356
	: : : : : :
Db	523 SPEPQYGGKDCVGDVTEHQMCNK RSCP IDGCLSNPCFP 560

RESULT 7

T18856

angiogenesis inhibitor homolog - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C;Accession: T18856; T24653

R;McMurray, A.

submitted to the EMBL Data Library, July 1995

A;Reference number: Z19031

A;Accession: T18856

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1444 <WIL>

A;Cross-references: EMBL:Z50004; PIDN:CAA90293.1; GSPDB:GN00028; CESP:C02B4.1

A;Experimental source: clone C02B4

R;McMurray, A.

submitted to the EMBL Data Library, July 1995

A;Reference number: Z19917

A;Accession: T24653

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1444 <WI2>

A;Cross-references: EMBL:Z50006; PIDN:CAA90302.1; GSPDB:GN00028; CESP:C02B4.1

A;Experimental source: clone T07C5

C;Genetics:

A;Gene: CESP:C02B4.1

A;Map position: X

A;Introns: 25/3; 70/3; 96/3; 139/3; 187/1; 234/2; 282/3; 376/2; 422/2; 478/3; 509/3; 566/2; 625/1; 696/2; 786/3; 812/2; 878/3; 971/1; 1007/3; 1067/1; 1099/3; 1180/3; 1273/2; 1305/1; 1363/1; 1388/2

Query Match 5.8%; Score 276; DB 2; Length 1444;
 Best Local Similarity 27.2%; Pred. No. 1.7e-11;
 Matches 73; Conservative 28; Mismatches 97; Indels 70; Gaps 12;

Qy	123 WCQCVAWSSSGTTKSQKAYIRIARLRKNFEQ-----EPLAKEVSLEQGIVLPCR PPEG I 176
	: : : : : : : :
Db	1134 WSEWSSWSAC-----SCFSLTSTRRRFCQVVDPTVQGFCAGAILEQ---IPCAPGSCS 1183

Qy	177 PPAE-----VEW-----LRNEDLVDP SLDP NVYITREHSLVVRQARLADTAN 218
----	--

Db	1184 PSAGGWSLWSEWSSCSKDCGDTGHQIRNRM CSEP-----IPSNRGAYCSG 1228
Qy	219 YT-----CVAKNIVARRRSASA AAVIVYVNGGWSTWTEWSVC SASCGRGWQKRSRSCTNPA 273 : : : : : : : :
Db	1229 YSF DQRPCVMDNVC SDEK-----VDGGWTDWTAWSECTDYCRNGHRSRTRFCANPK 1279
Qy	274 PLNGGAFC EGVQNVQKTAC--ATLC PV-DGSWSPWSKWSACGLDC---THWRSRECSDPAP 327 : : : :
Db	1280 PSQGGAQCTGSDFELNPCFD PARCHLRDGGWSTWSDWTPCSAS CGFGVQTRDRSCSSPEP 1339
Qy	328 RNGGEECQGT DLDTRNCTSDL CVHSASG 355 : : :
Db	1340 K-GGQSCSGLAHQTSLCDLPACDHESDG 1366

RESULT 8

T00326

hypothetical protein KIAA0550 - human

C;Species: Homo sapiens (man)

C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000

C;Accession: T00326

R;Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.

DNA Res. 5, 31-39, 1998

A;Title: Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.

A;Reference number: Z14086; MUID:98290545; PMID:9628581

A;Accession: T00326

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-984 <NAG>

A;Cross-references: EMBL:AB011122; NID:g3043623; PIDN:BAA25476.1; PID:g3043624

A;Experimental source: brain

C;Genetics:

A;Note: KIAA0550

C;Superfamily: thrombospondin type 1 repeat homology

F;344-398/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 5.7%; Score 275; DB 2; Length 984;
 Best Local Similarity 39.0%; Pred. No. 1.2e-11;
 Matches 57; Conservative 20; Mismatches 53; Indels 16; Gaps 6;

Qy	220 TCVA-----KNIVARRRSASA AAVIVYVNGGWSTWTEWSVC SASCGRGWQKRSRSCTNPA 273 : : : : : : :
Db	317 TCVSPYGTHCSGPLRESRVCNN TALCPVHG VWEWS PWLS LCSFTCGRGQRT RTR SCT--P 374
Qy	274 PLNGGAFC EGVQNVQKTAC-ATLC PV-DGSWSPWSKWSACGLDC---THWRSRECSDPAPRN 329 : :
Db	375 PQYGGRPCEGPETHHKPCNIALCPV DGQWQEWSWSQCSVTCSNGTQQRSRQCT--AAAH 432
Qy	330 GGEECQGT DLDTRNCTSDL CVHSASG 355 : : : :
Db	433 GGSEC RGWP AESRECYNPEC--TANG 456

RESULT 9
 T00028
 brain-specific angiogenesis inhibitor 3 - human
 N;Alternate names: BAI3 protein
 C;Species: Homo sapiens (man)
 C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
 C;Accession: T00028
 R;Shiratsuchi, T.; Nishimori, H.; Ichise, H.; Nakamura, Y.; Tokino, T.
Cytogenet. Cell Genet. 79, 103-108, 1997
 A;Title: Cloning and characterization of BAI2 and BAI3, novel genes homologous to brain-specific angiogenesis inhibitor 1 (BAI 1).
 A;Reference number: Z14066; MUID:98194217; PMID:9533023
 A;Accession: T00028
 A;Status: translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Residues: 1-1522 <SHI>
 A;Cross-references: EMBL:AB005299; NID:g3021700; PIDN:BAA25363.1; PID:g3021701
 A;Experimental source: brain
 C;Genetics:
 A;Gene: GDB:BAI3
 A;Cross-references: GDB:9838090; OMIM:602684
 A;Map position: 6q12-6q12
 C;Superfamily: thrombospondin type 1 repeat homology
 F;344-398/Domain: thrombospondin type 1 repeat homology <THR3>

```

Query Match          5.7%;  Score 275;  DB 2;  Length 1522;
Best Local Similarity 39.0%;  Pred. No. 2.1e-11;
Matches 57;  Conservative 20;  Mismatches 53;  Indels 16;  Gaps 6;
  
```

QY	220	TCVA-----KNIVARRRSASAIVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPA	273
	: : : : : : : : : :		
Db	317	TCVSPYGTHCSGPLRESRVCNNTALCPVHGVWEWSPWSLCSFTCGRGQRTTRRSCT--P	374
QY	274	PLNGGAFCEGQNQKTAC-ATLCPVDGWSPWSKWSACGLDC---THWRSRECSDPAPRN	329
	: : : :		
Db	375	PQYGGRPCEGPETHHKPCNIALCPVDGQWQEWSWSQCSVTCNSNGTQQRSRQCT--AAAH	432
QY	330	GGEECQGTLDLTRNCTS DLCVHSASG	355
	: :: .: :		
Db	433	GGSECRGPWAESRECYNPEC--TANG	456

RESULT 10
 T00027
 brain-specific angiogenesis inhibitor 2 - human
 N;Alternate names: BAI2 protein
 C;Species: Homo sapiens (man)
 C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
 C;Accession: T00027
 R;Shiratsuchi, T.; Nishimori, H.; Ichise, H.; Nakamura, Y.; Tokino, T.
Cytogenet. Cell Genet. 79, 103-108, 1997
 A;Title: Cloning and characterization of BAI2 and BAI3, novel genes homologous to brain-specific angiogenesis inhibitor 1 (BAI 1).
 A;Reference number: Z14066; MUID:98194217; PMID:9533023
 A;Accession: T00027
 A;Status: translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA

A;Residues: 1-1572 <SHI>
A;Cross-references: EMBL:AB005298; NID:g3021698; PIDN:BAA25362.1; PID:g3021699
A;Experimental source: brain
C;Genetics:
A;Gene: GDB:BAI2
A;Cross-references: GDB:9838089; OMIM:602683
A;Map position: 1p35-1p35

Query Match 5.7%; Score 274.5; DB 2; Length 1572;
Best Local Similarity 19.2%; Pred. No. 2.4e-11;
Matches 176; Conservative 108; Mismatches 307; Indels 327; Gaps 38;

Qy 173 PEGIPPAEVEWLRNEDLVDPSDLDPNVY-----ITREHSLVVRQARL 213
|| | : | : | : | : | : | : | || | | |
Db 271 PEEEPKVKTQWPRSAD----EPGLYMAQTGDPAAEWSPWSVCSTCGQGLQVR-TRS 323

Qy 214 ADTANYTCVAKNIVARRRSASAIVIVYVNGWSTWTEWSVCSASCGRGWQKRSRSCTNPA 273
:: | : : | : : | : | : | | || : | : | : | : |
Db 324 CVSSPYGTLCSGPLRETRPCNNSATCPVHGWEWGSWSLCSRSCRGSRSRMRTCV--P 381

Qy 274 PLNGGAFCCEGQNVQKTACA-TLCPVDGWSPWSKWSACGLDC---THWRSRECSDPAPR- 328
| : || | || : | : | || | | | | | | | | | | |
Db 382 PQHGGKACEGPELQTKLCSMAACPVEGQWLEWGPWGPCSTSCANGTQQRSRKCSVAGPAW 441

Qy 329 -----NGGEECQ 335
| : |
Db 442 ATCTGALTDTRECSNLECPATDSKWGPWNAWSLCSKTCDTGWQRRFRMCQATGTQGYCPE 501

Qy 336 GTLDLDRNCTS DLC--VHSASGPEDVAL----- 361
| | : : | : | | | | | |
Db 502 GTGEEVKPCSEKRCPAFHEMCRDEYVMLMTWKAAAGEIIYNKCPPNASGSASRRCLLSA 561

Qy 362 ---YVGLIAVAVCL---VLLLLVLILVYCRKKEGLSDVADSSIITSGFQPVSIKPSKA 414
| | | : | | : | | : | : | : | : | : | : | : |
Db 562 QGVAYWGLPSFARCISHEYRYLYSLREHLAKGQRMLAGEGMSQVVR-S-LQELLARRYY 620

Qy 415 DNPHLLTIQPDLS TTTTYQGS LCP RQDG PSP KFQLT-----NGHLLSPLGG 461
| : : | | : | | | | | | | | | | | | | | | |
Db 621 SGD LLS FVD I L R N V T D T F K R A T Y V P S A D D V Q R F F Q V V S F M V D A E N K E K W D D A Q Q V S P --G 678

Qy 462 GRHTLHHSSPTSEAEFFV-----SRLSTQNYFRSLPRG-----TSNMTYGTFN 504
| | | | | | | | | | | | | | | | | | | | | | | |
Db 679 SVHLLR-----VVEDFIHLVGDA LKA FQSSLIVTDNLVISIQREP VSAVSSDIT FPMRG 732

Qy 505 FLG-----GRLMIPNTG I S L L I P-----PDAI PRG K----- 530
| | | | | | | | | | | | | | | | | | | | | | | |
Db 733 RRGMKD WVRH SED RLFLPKEV L S LSSPGKPATSGAAGSPGRGRGPGTVPPGPGHSHQ RLL 792

Qy 531 -----IYE-IYLTLHKPEDVRLPLAGCQ TL SPIVSCGPPGVLLTRPVIL 574
| | : | | | | | | | | | | | | | | | | | | | | |
Db 793 PADPDESSYFVIGAVLYRTLGLL PPP--RPPLAVTSRVMT--VTVRPPTQOPPAEPLIT 847

Qy 575 A-----MDHC GE P S P D S W S L R L K K Q S C E G S W E D V L H L G E E A P S H L Y Y C Q -LEASAC Y V-- 626
| | : | | | | | | | | | | | | | | | | | | | | | |
Db 848 VEL SYI ING TT DPHC ASWDYS-RADAS SGD WD-----TENC QT LET QAAH TRC 894

Qy 627 FTEQLGRFALVGE-----ALSVA AKRLKLLL FAPVACTSLEYNI RVY CLHD TH DALKEV 681

		: :: : : : : : : :
Db	895	QCQHLSTFAVLAQPPKDLTLEAGSPSVPLVIGCAVSCMALLTLLAIYA-----AFWRF 948
Qy	682	VQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVSSLWKSKLLVSYQEIPFYHIWNGTQ 741
		:: : : :: : : :
Db	949	IKSERSI-----ILLNFCLSI--LASNI---LILVGQSRVLSKGVCTMTA 988
Qy	742	RYLHCTFTLERVSPSTSDLACKLWV-----WQVEGDG 773
		: : : :
Db	989	AFLHFFF-----LSSFCWLTEAWQSYLAVIGRMTRLVRKRFLCLGWGLPALV 1037
Qy	774	QSFSINFNITKDTRFAELLALESEAG-VPALVGPSA-----FKIPFLIRQKI----IS 821
		: : : : :
Db	1038	VAVSVGFRTKGYGTSSYCWLSSLEGGLLYAFVGPAAVIVLVNMLIGIIVFNKLMARDGIS 1097
Qy	822	SLDPPCRRGAD---WRTL 836
		:: :
Db	1098	DKSKKQRAGSERCPWASL 1115

RESULT 11

A40558

thrombospondin 1 precursor - mouse

C;Species: Mus musculus (house mouse)

C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 20-Aug-1999

C;Accession: A40558; A37905; B42587; S68787

R;Lawler, J.; Duquette, M.; Ferro, P.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.

Genomics 11, 587-600, 1991

A;Title: Characterization of the murine thrombospondin gene.

A;Reference number: A40558; MUID:92128941; PMID:1774063

A;Accession: A40558

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1170 <LAW>

A;Cross-references: GB:M62449; GB:M62450; GB:M62451; GB:M62452; GB:M62453; GB:M62454; GB:M62455; GB:M62456; GB:M62457; GB:M62458; GB:M62459; GB:M62460; GB:M62461; GB:M62462; GB:M62463; GB:M62464; GB:M62465; GB:M62466; GB:M62467; GB:M62468; GB:M62469; GB:M62470; NID:g511867; PIDN:AAA50611.1; PID:g511869

R;Bornstein, P.; Alfi, D.; Devarayalu, S.; Framson, P.; Li, P.

J. Biol. Chem. 265, 16691-16698, 1990

A;Title: Characterization of the mouse thrombospondin gene and evaluation of the role of the first intron in human gene expression.

A;Reference number: A37905; MUID:90375546; PMID:2398070

A;Accession: A37905

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-490 <BOR>

A;Cross-references: GB:J05605; GB:J05606; NID:g201991; PIDN:AAA40431.1; PID:g554390

R;Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M. J. Biol. Chem. 267, 3274-3281, 1992

A;Title: Characterization of mouse thrombospondin 2 sequence and expression during cell growth and development.

A;Reference number: A42587; MUID:92147683; PMID:1371115

A;Accession: B42587

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA
A;Residues: 1-1152, 'P', 1154-1170 <LAH>
A;Cross-references: GB:M87276
A;Note: sequence extracted from NCBI backbone (NCBIP:81501)
R;Chen, H.; Aeschlimann, D.; Nowlen, J.; Mosher, D.F.
FEBS Lett. 387, 36-41, 1996
A;Title: Expression and initial characterization of recombinant mouse thrombospondin 1 and thrombospondin 3.
A;Reference number: S68787; MUID:96234006; PMID:8654563
A;Accession: S68787
A;Molecule type: protein
A;Residues: 19-26, 'X', 28-37 <CHE>
C;Complex: homotrimer, disulfide linked
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology
C;Keywords: calcium binding; glycoprotein; homotrimer
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-1170/Product: thrombospondin 1 #status predicted <MAT>
F;317-375/Domain: von Willebrand factor type C repeat homology <VWC>
F;378-429/Domain: thrombospondin type 1 repeat homology <THR1>
F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>
F;491-547/Domain: thrombospondin type 1 repeat homology <THR3>
F;551-586/Domain: EGF homology <EGF>
F;248,360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.6%; Score 270.5; DB 2; Length 1170;
Best Local Similarity 32.2%; Pred. No. 3.2e-11;
Matches 57; Conservative 24; Mismatches 71; Indels 25; Gaps 5;

Qy	207 VVRQARLADTANYTCVAKNIVAR-----RRSASAIVIVYVNGGWSTWTEWSVCAS 258
	: :: : :: : : : :
Db	399 IQQQRGRSCDSLNNRCEGSSVQRTCHIQECDKRFKQ-----DGGWSHWSPWSSCSVTC 451
Qy	259 GRGWQKRSRSCTNPAPLNGGAFCEGQNQKTAC-ATLCPVDGWSWPSKWSACGLDC--- 314
	: : : : : :
Db	452 GDGVITRIRLCNSPSPQMNGKPCEGEARETKACKDACPINGGWGPWSPWDICSVTCGGG 511
Qy	315 THWRSRECSDPAPRNGGEECQGTLDTRNCTSDLCVHSASGPEDVALYVGLIAAVC 371
	: : : :
Db	512 VQRRSRRLCNPNTPQFGGKDCVGDVTENQVCNKQDC-----PIDGCLSNPCFAGAKC 562

RESULT 12
TSHUP1
thrombospondin 1 precursor - human
C;Species: Homo sapiens (man)
C;Date: 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change 17-Nov-2000
C;Accession: A26155; A34274; A30140; A25812; A05172; A42927
R;Lawler, J.; Hynes, R.O.
J. Cell Biol. 103, 1635-1648, 1986
A;Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple calcium-binding sites and homologies with several different proteins.
A;Reference number: A26155; MUID:87057617; PMID:2430973
A;Accession: A26155
A;Molecule type: mRNA
A;Residues: 1-1170 <LAH>
A;Cross-references: GB:X04665; NID:g37137; PIDN:CAA28370.1; PID:g37138

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing
R;Laherty, C.D.; Gierman, T.M.; Dixit, V.M.
J. Biol. Chem. 264, 11222-11227, 1989
A;Title: Characterization of the promoter region of the human thrombospondin gene. DNA sequences within the first intron increase transcription.
A;Reference number: A34274; MUID:89291870; PMID:2544587
A;Accession: A34274
A;Molecule type: DNA
A;Residues: 1-166 <LAH>
A;Cross-references: GB:J04835
R;Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, P.; Frazier, W.A.
J. Cell Biol. 108, 729-736, 1989
A;Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the 3' untranslated region.
A;Reference number: A30140; MUID:89139590; PMID:2918029
A;Accession: A30140
A;Molecule type: mRNA
A;Residues: 1-83,'A',85-522,'A',524-1170 <HEN>
A;Cross-references: EMBL:X14787; NID:g37464; PIDN:CAA32889.1; PID:g37465
A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing
R;Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.
Biochemistry 25, 8418-8425, 1986
A;Title: Partial amino acid sequence of human thrombospondin as determined by analysis of cDNA clones: homology to malarial circumsporozoite proteins.
A;Reference number: A25812; MUID:87157592; PMID:3030396
A;Accession: A25812
A;Molecule type: mRNA
A;Residues: 1-83,'A',85-397 <KOB>
A;Cross-references: GB:M25631; NID:g538353; PIDN:AAA36741.1; PID:g538354
R;Dixit, V.M.; Hennessy, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986
A;Reference number: A05172; MUID:86287276; PMID:3461443
A;Accession: A05172
A;Molecule type: mRNA
A;Residues: 1-83,'A',85-374,'RC' <DIX>
A;Cross-references: GB:M14326; NID:g340005; PIDN:AAA61237.1; PID:g553801
A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing
R;Sun, X.; Skorstengaard, K.; Mosher, D.F.
J. Cell Biol. 118, 693-701, 1992
A;Title: Disulfides modulate RGD-inhibitable cell adhesive activity of thrombospondin.
A;Reference number: A42927; MUID:92348511; PMID:1379247
A;Accession: A42927
A;Molecule type: protein
A;Residues: 987-1003 <SUN>
A;Note: Cys-992 is shown to have a free sulphydryl
C;Genetics:
A;Gene: GDB:THBS1; TSPL; TSP
A;Cross-references: GDB:120438; OMIM:188060
A;Map position: 15q15-15q15
A;Introns: 23/1
A;Note: the list of introns may be incomplete
C;Complex: homotrimer, disulfide linked

C;Function:
A;Description: participates in cell migration and adhesion, and in platelet aggregation
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology
C;Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-1170/Product: thrombospondin 1 #status predicted <MAT>
F;317-375/Domain: von Willebrand factor type C repeat homology <VWC>
F;378-429/Domain: thrombospondin type 1 repeat homology <THR1>
F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>
F;491-547/Domain: thrombospondin type 1 repeat homology <THR3>
F;551-586/Domain: EGF homology <EGF1>
F;650-689/Domain: EGF homology <EGF2>
F;926-928/Region: cell attachment (R-G-D) motif
F;171-232/Disulfide bonds: #status predicted
F;248,360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;270,274/Disulfide bonds: interchain #status predicted
F;610/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F;1051/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 5.6%; Score 268.5; DB 1; Length 1170;
Best Local Similarity 32.9%; Pred. No. 4.4e-11;
Matches 51; Conservative 24; Mismatches 61; Indels 19; Gaps 4;

Qy	207 VVRQARLADTANYTCVAKNIVAR-----RRSASAAVIVYVNGGWSTWTEWSVCSASC 258
	: :: : :: : : : :
Db	399 IQQQRGRSCDSLNNRCEGSSVQTRTCHIQECDKRFKQ-----DGGWSHWSPWSSCSVTC 451
Qy	259 GRGWQKRSRSCTNPAPLNNGAFCEGQNVQKTAC-ATLCPVDGWSWPWSKWSACGLDC--- 314
	: : : : :
Db	452 GDGVITRIRLCNSPSPQMNGKPCEGEARETKACKDACPINGGWGPWSPWDICSVTCGGG 511
Qy	315 THWRSRECSDPAPRNGGEECQGTLDTRNCTS DLC 349
	: : : : :
Db	512 VQKRSRLCNNPTPQFGGKDCVGDTVENQICNKQDC 546

RESULT 13

A39804
thrombospondin precursor - chicken
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A39804
R;Lawler, J.; Duquette, M.; Ferro, P.
J. Biol. Chem. 266, 8039-8043, 1991
A;Title: Cloning and sequencing of chicken thrombospondin.
A;Reference number: A39804; MUID:91217026; PMID:2022631
A;Accession: A39804
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1178 <LAW>
A;Cross-references: GB:M60853; NID:g212763; PIDN:AAA51437.1; PID:g212764
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology
F;325-383/Domain: von Willebrand factor type C repeat homology <VWC>

F;386-437/Domain: thrombospondin type 1 repeat homology <THR1>
F;442-498/Domain: thrombospondin type 1 repeat homology <THR2>
F;499-555/Domain: thrombospondin type 1 repeat homology <THR3>
F;658-697/Domain: EGF homology <EGF>

RESULT 14

S29126
properdin precursor [validated] - human
N;Alternate names: factor P
C;Species: Homo sapiens (man)
C;Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
C;Accession: S29126; S16150; A05319; T45112; T45113
R;Nolan, K.F.; Kaluz, S.; Higgins, J.M.G.; Goundis, D.; Reid, K.B.M.
Biochem. J. 287, 291-297, 1992
A;Title: Characterization of the human properdin gene.
A;Reference number: S29126; MUID:93038568; PMID:1417780
A;Accession: S29126
A;Molecule type: DNA
A;Residues: 1-469 <NOL1>
A;Cross-references: EMBL:X70872; NID:g35679; PIDN:CAA50220.1; PID:g35680
R;Nolan, K.F.; Schwaeble, W.; Kaluz, S.; Dierich, M.P.; Reid, K.B.M.
Eur. J. Immunol. 21, 771-776, 1991
A;Title: Molecular cloning of the cDNA coding for properdin, a positive
regulator of the alternative pathway of human complement.
A;Reference number: S16150; MUID:91184288; PMID:2009915
A;Accession: S16150
A;Molecule type: mRNA
A;Residues: 1-456,'R',458-469 <NOL2>
A;Cross-references: EMBL:X57748
R;Reid, K.B.M.; Gagnon, J.
Mol. Immunol. 18, 949-959, 1981
A;Reference number: A05319; MUID:82195224; PMID:7341961
A;Accession: A05319
A;Molecule type: protein
A;Residues: 28-53,'Q',55-59,'G',61,'I',63;137-138,'P',140-141,'P',143-
144,'X',146-148,'Y',150,'S',152,'Y',154-156,'XSXGXAX';162-163,'E',165-
172,'X',174-176,'X',178,'V',180;223-228,'X',230-232,'GX',235-238,'GH',241-
245;248-251,'X',253-257,'P',259,'G',261,'XPP',265-266,'X',268-269;280-
285,'X',287-290,'X',292,'H',294-300,'SXXXX',305-307,'X',309-315,'K',317;333-

341,343-357,'X',359-362,'EXE';393-404,'QK',407;421-427,'R',429-443,'TKV',447-448,'XX',451,'RX',454-455 <REI>
R;Westberg, J.; Nordin-Fredrikson, G.; Truedsson, L.; Sjoholm, A.G.; Uhlen, M.
submitted to the EMBL Data Library, May 1997
A;Reference number: Z22914
A;Accession: T45112
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-54,'X',56-73,'X',75-99,'W',101-469 <WES1>
A;Cross-references: EMBL:AF005665; PIDN:AAB63280.1
A;Experimental source: genomic DNA from individual with properdin deficiency
type II
A;Accession: T45113
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-60,'X',62-413,'D',415-452,'XX',455-469 <WE2>
A;Cross-references: EMBL:AF005666; PIDN:AAC51626.1
A;Experimental source: genomic DNA from individual with properdin deficiency
type III
R;Hartmann, S.; Hofsteenge, J.
J. Biol. Chem. 275, 28569-28574, 2000
A;Title: Properdin, the positive regulator of complement, is highly C-
mannosylated.
A;Reference number: A59360; MUID:20435812; PMID:10878002
A;Contents: annotation
A;Note: identification and location of C-mannosylation sites by mass-
spectroscopy
C;Genetics:
A;Gene: GDB:PFC
A;Cross-references: GDB:120275; OMIM:312060
A;Map position: Xp11.3-Xp11.23
A;Introns: 26/1; 76/2; 135/1; 192/1; 256/1; 314/1; 378/1; 415/2
C;Complex: a mixture of homodimers, homotrimers and homotetramers
C;Function:
A;Description: protects C3 convertase (C3bBb) from rapid inactivation
A;Pathway: complement alternate pathway
C;Superfamily: human properdin precursor; thrombospondin type 1 repeat homology
C;Keywords: complement alternate pathway; glycoprotein; homodimer; homotetramer;
homotrimer; plasma
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-469/Product: properdin #status experimental <MAT>
F;76-128/Domain: thrombospondin type 1 repeat homology <THR1>
F;135-191/Domain: thrombospondin type 1 repeat homology <THR2>
F;192-255/Domain: thrombospondin type 1 repeat homology <THR3>
F;256-313/Domain: thrombospondin type 1 repeat homology <THR4>
F;314-377/Domain: thrombospondin type 1 repeat homology <THR5>
F;378-440/Domain: thrombospondin type 1 repeat homology <THR6>
F;83,86,139,142,145,196,199,260,263,321,324,382,385,388/Modified site: 2'-
mannosyl-tryptophan (Trp) #status experimental
F;428/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.1%; Score 243; DB 1; Length 469;
Best Local Similarity 39.5%; Pred. No. 9.2e-10;
Matches 45; Conservative 14; Mismatches 43; Indels 12; Gaps 4;

QY 243 GGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACAT--LCPVDGS 300
|||| | | || :| :| :| | :|| | | | | :| | | :|| | :

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2004, 14:33:31 ; Search time 56 Seconds
(without alignments)
4991.658 Million cell updates/sec

Title: US-10-624-932-2

Perfect score: 4791

Sequence: 1 MAVRPGLWPALLGIVLAAWL.....AVAGLGQPDAGLFTVSEAEC 898

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1276540 seqs, 311283816 residues

Total number of hits satisfying chosen parameters: 1276540

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*

10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*

11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*

13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*

14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*

15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*

17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%

Result	Query	No.	Score	Match	Length	DB	ID	Description
--------	-------	-----	-------	-------	--------	----	----	-------------

1	4791	100.0	898	12	US-09-918-779-2	Sequence 2, Appli
2	4791	100.0	898	16	US-10-624-932-2	Sequence 2, Appli
3	4698.5	98.1	899	11	US-09-970-944-2	Sequence 2, Appli
4	4638	96.8	898	10	US-09-933-261-5	Sequence 5, Appli
5	4638	96.8	898	11	US-09-970-944-13	Sequence 13, Appli
6	4638	96.8	898	14	US-10-256-702-5	Sequence 5, Appli
7	4638	96.8	898	14	US-10-240-154-16	Sequence 16, Appli
8	4413	92.1	842	16	US-10-311-623-1	Sequence 1, Appli
9	2845	59.4	544	11	US-09-970-944-14	Sequence 14, Appli
10	2815.5	58.8	557	10	US-09-933-261-6	Sequence 6, Appli
11	2815.5	58.8	557	14	US-10-256-702-6	Sequence 6, Appli
12	2787	58.2	931	11	US-09-970-944-15	Sequence 15, Appli
13	2787	58.2	931	12	US-10-087-684-35	Sequence 35, Appli
14	2787	58.2	931	12	US-09-972-211-121	Sequence 121, App
15	2787	58.2	931	12	US-10-037-417-117	Sequence 117, App
16	2787	58.2	931	12	US-10-096-625-121	Sequence 121, App
17	2787	58.2	1010	12	US-10-218-779-35	Sequence 35, Appli
18	2762	57.6	931	11	US-09-970-944-16	Sequence 16, Appli
19	2762	57.6	931	12	US-09-972-211-125	Sequence 125, App
20	2762	57.6	931	12	US-10-096-625-125	Sequence 125, App
21	2759	57.6	931	12	US-10-037-417-120	Sequence 120, App
22	2755	57.5	931	11	US-09-970-944-17	Sequence 17, Appli
23	2755	57.5	931	12	US-10-087-684-36	Sequence 36, Appli
24	2755	57.5	931	12	US-10-218-779-36	Sequence 36, Appli
25	2755	57.5	931	12	US-09-972-211-122	Sequence 122, App
26	2755	57.5	931	12	US-10-037-417-118	Sequence 118, App
27	2755	57.5	931	12	US-10-037-417-119	Sequence 119, App
28	2755	57.5	931	12	US-10-096-625-122	Sequence 122, App
29	2578.5	53.8	945	12	US-10-087-684-34	Sequence 34, Appli
30	2578.5	53.8	945	12	US-10-218-779-34	Sequence 34, Appli
31	2578.5	53.8	945	12	US-09-972-211-124	Sequence 124, App
32	2578.5	53.8	945	12	US-10-037-417-121	Sequence 121, App
33	2578.5	53.8	945	12	US-10-096-625-124	Sequence 124, App
34	2572.5	53.7	945	12	US-10-087-684-33	Sequence 33, Appli
35	2572.5	53.7	945	12	US-10-218-779-33	Sequence 33, Appli
36	2572.5	53.7	945	12	US-09-972-211-123	Sequence 123, App
37	2572.5	53.7	945	12	US-10-096-625-123	Sequence 123, App
38	2571.5	53.7	943	10	US-09-933-261-7	Sequence 7, Appli
39	2571.5	53.7	943	14	US-10-256-702-7	Sequence 7, Appli
40	2563.5	53.5	933	12	US-10-087-684-2	Sequence 2, Appli
41	2563.5	53.5	933	12	US-10-087-684-4	Sequence 4, Appli
42	2563.5	53.5	933	12	US-10-218-779-2	Sequence 2, Appli
43	2563.5	53.5	933	12	US-10-218-779-4	Sequence 4, Appli
44	2558.5	53.4	945	12	US-10-147-493-146	Sequence 146, App
45	2558.5	53.4	945	12	US-10-145-127-146	Sequence 146, App

)

ALIGNMENTS

RESULT 1

US-09-918-779-2

; Sequence 2, Application US/09918779

; Publication No. US20030064369A1

; GENERAL INFORMATION:

; APPLICANT: Taupier, Raymond

; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Rastelli, Luca
; APPLICANT: Spaderna, Steven
; APPLICANT: Shimkets, Richard
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Spytek, Kimberly
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Gusev, Vladimir
; APPLICANT: Grosse, William
; APPLICANT: Alsobrook, John
; APPLICANT: Lepley, Denise
; APPLICANT: Burgess, Catherine
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Stone, David
; APPLICANT: Smithson, Glennda
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-074 US
; CURRENT APPLICATION NUMBER: US/09/918,779
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/221,409
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/222,840
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,752
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,762
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,770
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,769
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/225,146
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,392
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/225,470
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/225,697
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: 60/263,662
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: 60/281,645
; PRIOR FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-918-779-2

Query Match 100.0%; Score 4791; DB 12; Length 898;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 898; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
|||
Db 1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy 61 VLLVCKAVPATQIFFKCNGEWRQVDHIERSTDGSSGLPTMEVRINVSQQVEKVFGL 120
|||
Db 61 VLLVCKAVPATQIFFKCNGEWRQVDHIERSTDGSSGLPTMEVRINVSQQVEKVFGL 120

Qy 121 EYWQCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
|||
Db 121 EYWQCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy 181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVY 240
|||
Db 181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVY 240

Qy 241 VNGGWSTWTEWSVCASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGS 300
|||
Db 241 VNGGWSTWTEWSVCASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGS 300

Qy 301 WSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDRNCTS DLCVHSASGPEDVA 360
|||
Db 301 WSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDRNCTS DLCVHSASGPEDVA 360

Qy 361 LYVGLIAAVAVCLVLLLLVLILVYCRKKEGLSDVADSSI LTSGFQPVSIKPSKADNP HLL 420
|||
Db 361 LYVGLIAAVAVCLVLLLLVLILVYCRKKEGLSDVADSSI LTSGFQPVSIKPSKADNP HLL 420

Qy 421 TIQPDLS TTTTYQGS LCP RDGP SPKF QLT NGHLLS PLGG RHT LHSS PTSEAE FVS 480
|||
Db 421 TIQPDLS TTTTYQGS LCP RDGP SPKF QLT NGHLLS PLGG RHT LHSS PTSEAE FVS 480

Qy 481 RLSTQNYFRSLPRGTSNM TYGTFN FLG GR LMIPNT G ISL I PPD AIPRG KIYEI YLT LHK 540
|||
Db 481 RLSTQNYFRSLPRGTSNM TYGTFN FLG GR LMIPNT G ISL I PPD AIPRG KIYEI YLT LHK 540

Qy 541 PEDVRLPLAGC QTLLS PIVSCGPPGVLLTRPVILAMDHC GEPS PD SWSLR LKKQS CEG SW 600
|||
Db 541 PEDVRLPLAGC QTLLS PIVSCGPPGVLLTRPVILAMDHC GEPS PD SWSLR LKKQS CEG SW 600

Qy 601 EDVLHLGEEAPSHLYYCQ LEASAC YVFT EQLGRF ALVGEAL SVAA KRLK LLLF APVACT 660
|||
Db 601 EDVLHLGEEAPSHLYYCQ LEASAC YVFT EQLGRF ALVGEAL SVAA KRLK LLLF APVACT 660

Qy 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDV PSSLW 720
|||
Db 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDV PSSLW 720

Qy 721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS DLACKLWVWQVEGDGQSFSINF 780
|||
Db 721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS DLACKLWVWQVEGDGQSFSINF 780

Qy 781 NITKDTRFAELLALESEAGVPA LVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKL 840
|||
Db 781 NITKDTRFAELLALESEAGVPA LVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKL 840

Qy 841 HLDHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 898
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 841 HLDHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 898

RESULT 2

US-10-624-932-2

; Sequence 2, Application US/10624932
; Publication No. US20040096877A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Rastelli, Luca
; APPLICANT: Spaderna, Steven
; APPLICANT: Shimkets, Richard
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Spytek, Kimberly
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Gusev, Vladimir
; APPLICANT: Grosse, William
; APPLICANT: Alsobrook, John
; APPLICANT: Lepley, Denise
; APPLICANT: Burgess, Catherine
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Stone, David
; APPLICANT: Smithson, Glennda
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-074 US
; CURRENT APPLICATION NUMBER: US/10/624,932
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: 09/918,779
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/221,409
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/222,840
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,752
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,762
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,770
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,769
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/225,146
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,392
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/225,470
; PRIOR FILING DATE: 2000-08-15
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2

; LENGTH: 898
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-624-932-2

Query Match 100.0%; Score 4791; DB 16; Length 898;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 898; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVGANPDLLPHFLVEPEDVYIVKNKP 60
Db |||||||
Qy 61 VLLVCKAVPATQIFFKCNGEWVRQVDHIERSTDGSSGLPTMEVRINVSRRQQVEKVFGL 120
Db |||||||
Qy 121 EYWCQCVAWSSSGTTKSQKAYIRIARLRNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
Db |||||||
Qy 181 VEWLRNEDLVDPSDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVIVY 240
Db |||||||
Qy 241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCNPAPLNGGAFCEGQNVQKTACATLCPVDG 300
Db |||||||
Qy 301 WSPWSKWSACGLDCTHWRSRECSDPAPRNGGECCQGTLDTRNCTSDLVHSASGPEDVA 360
Db |||||||
Qy 361 LYVGLIAAVCLVLLLLVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPPLL 420
Db |||||||
Qy 421 TIQPDLSTTTTYQGSCLCPRQDGSPKFQLTNGHLLSPGGGRHTLHHSSPTSEAEFFVS 480
Db |||||||
Qy 481 RLSTQNYFRSLPRGTSNMTRYGTFNFLGGRLMPNTGISLLIPPDAIPRGKIYEIYLTLHK 540
Db |||||||
Qy 541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDSWSRLKKQSCEGSW 600
Db |||||||
Qy 601 EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLLFAFPACT 660
Db |||||||
Qy 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVSSLW 720
Db |||||||
Qy 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVSSLW 720

Qy	721	KSKLLVSYQEIPFYHIWNGTQRYLHCTTLERVSPSTSDLACKLWVWQVEGDGQSFSINF	780
Db	721	KSKLLVSYQEIPFYHIWNGTQRYLHCTTLERVSPSTSDLACKLWVWQVEGDGQSFSINF	780
Qy	781	NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKISSLDPCCRGA DWRTLAQKL	840
Db	781	NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKISSLDPCCRGA DWRTLAQKL	840
Qy	841	HLDSHLSFFASKPSP TAMILNLWEARHF PNGNLSQLAAAVAGLGQPDAGLFTVSEAEC	898
Db	841	HLDSHLSFFASKPSP TAMILNLWEARHF PNGNLSQLAAAVAGLGQPDAGLFTVSEAEC	898

RESULT 3

US-09-970-944-2

; Sequence 2, Application US/09970944

; Publication No. US20030204052A1

; GENERAL INFORMATION:

; APPLICANT: Herrman, John

; APPLICANT: Rastelli, Luca

; APPLICANT: Shimkets, Richard A

; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same and

; TITLE OF INVENTION: Antibodies Directed Against these Proteins

; FILE REFERENCE: 21402-138

; CURRENT APPLICATION NUMBER: US/09/970,944

; CURRENT FILING DATE: 2002-05-02

; PRIOR APPLICATION NUMBER: 60/237,862

; PRIOR FILING DATE: 2000-10-04

; NUMBER OF SEQ ID NOS: 62

SOFTWARE:

; SEQ ID NO 2

; LENGTH: 8
TYPE: BPT

ORGANISM: Homo sapiens

; ORGANISM: H
HS 08 870 844 3

03-09-970-944-2

Query Match 98.1% Score 4698.5 DB 11 Length 899:

Best Local Similarity 98.7%: Pred. No. 0;

Qy 1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 61 VLLVCKAVPATQI FFKCNGEVRQVDHVIERSTDGSSGLPTMEVRINVSROOQEKVFGLE 120

Db 61 VLLVCKAVPATQI FFKCNGEWWVRQVDHVI ERSTDGSSGEPTMEVRINVSQQVEKVFGL 120

121. ЕМІСОСУННІСС СТТКСОҚ НЕРЛІДЕ БҮНДЕСЕРІЛКЕНДІ ЛЕСІМІНДЕРЕДЕСІРІЛДЕ 122

DB 121 EYWQCVAVSSGTRSQKAYTRIARLRKFQEPLAKEVSLEQGIVLPCRPPEGIFPAE 180

181 VEWERNEDLVDFESLDENVITIREASLEVVRQARLADIANITCVARNIVARRSASAIVIVI 240

Pb 181 VEWLBNEDLVDPSTIDPNVYITREHSLIVRQABLAADTANYTCVAKNTIVARRBSASAATVY 240

Qy	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNV-QKTACATLCPVDG	299
			: :
Db	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVHDRTVSSLVSDG	300
Qy	300	SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGECCQGTDLDRNCTS DLCVHSASGPEDV	359
Db	301	SWSPWSKWSACGLDCTHWRSRECSDPAPRNGGECCQGTDLDRNCTS DLCVHSASGPEDV	360
Qy	360	ALYVGLIAAVCLVLLLLVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPHL	419
Db	361	ALYVGLIAAVCLVLLLLVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPHL	420
Qy	420	LTIQPDLSTTTTYQGSLCPRQDGSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFFV	479
Db	421	LTIQPDLSTTTTYQGSLCPRQDGSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFFV	479
Qy	480	SRLSTQNYFRSLPRGTSNMTYGTNFNLGGRLMIPNTGISLLIPPPDAIPRGKIYEIYLTLH	539
Db	480	SRLSTQNYFRSLPRGTSNMTYGTNFNLGGRLMIPNTGISLLIPPPDAIPRGKIYEIYLTLH	539
Qy	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGS	599
Db	540	KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGS	599
Qy	600	WE-DVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSAAKRLKLLLFA PVA	658
Db	600	WEQDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSAAKRLKLLLFA PVA	659
Qy	659	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKD SYHNLRLSIHDV PSS	718
Db	660	CTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKD SYHNLRLSIHDV PSS	719
Qy	719	LWKS KL LVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS DLACKLWVWQVEGDQSF SI	778
Db	720	LWKS KL LVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS DLACKLWVWQVEGDQSF SI	779
Qy	779	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ	838
Db	780	NFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ	839
Qy	839	KLHLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAA VAGLGQPDAGLFTVSEAEC	898
Db	840	KLHLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAA VAGLGQPDAGLFTVSEAEC	899

RESULT 4

US-09-933-261-5

; Sequence 5, Application US/09933261

; Publication No. US20030040046A1

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/933,261
; FILING DATE: 20-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/808,982
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 898 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20030040046A1 Relevant
; TOPOLOGY: No. US20030040046A1 Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-933-261-5

Query Match 96.8%; Score 4638; DB 10; Length 898;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 862; Conservative 17; Mismatches 19; Indels 0; Gaps 0;

Qy	1 MAVRPGLWPALLGIVLAALRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
Db	1 MAVRPGLWPVLLGIVLAALRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
Qy	61 VLLVCKAVPATQIFFKCNGEVRQVDHVIERSTDGSSGLPTMEVRINVSQQVEKVGLE 120
Db	61 VLLVCKAVPATQIFFKCNGEVRQVDHVIERSTDSSGLPTMEVRINVSQQVEKVGLE 120
Qy	121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
Db	121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
Qy	181 VEWLRNEDLVDPSDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVIVY 240
Db	181 VEWLRNEDLVDPSDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSTSAAIVIVY 240
Qy	241 VNNGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGS 300

Db	241	VNGGWSTWESVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDG	300
Qy	301	WSPWSKWSACGLDCTHWRSRECSDPAPRNGGECCQGTDLDTRNCTS DLCVHSASGPEDVA	360
Db	301	WSSWSKWSACGLDCTHWRSRECSDPAPRNGGECCRGA LDTRNCTS DLCLHTASCPEDVA	360
Qy	361	LYVGLIAVAVCLLLLLVLTLVYCRKKEGLSDVADSSILTSGFQPVS IKPSKADNPHLL	420
Db	361	LYIGLVA VAVCLFL LLALGLI YCRKKEGLSDVADSSILTSGFQPVS IKPSKADNPHLL	420
Qy	421	TIQPDLS TTTTYQGS LCPQDGSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFVS	480
Db	421	TIQPDLS TTTTYQGS LCSRQDGSPKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEDFVS	480
Qy	481	RLSTQNYFRSLPRGT SNMTYGT FNFLGGRLMIPNTGISLLIPPDAI PRGK IYEIYLTLHK	540
Db	481	RLSTQNYFRSLPRGT SNMAYGT FNFLGGRLMIPNTGISLLIPPDAI PRGK IYEIYLTLHK	540
Qy	541	PEDVRLPLAGCQTLLS PIVSCGPPGVLLTRPVILAMDHC GEPSPDSW SLRLKKQSCEGSW	600
Db	541	PEDVRLPLAGCQTLLSPV VSCGPPGVLLTRPVILAMDHC GEPSPDSW SLRLKKQSCEGSW	600
Qy	601	EDVLHLGEEAPSH LYCCQLEASAC YVFTEQLGRFALVGEALSVA AKRLKLLL FAPVACT	660
Db	601	EDVLHLGEEPSH LYCCQLEAGAC YVFTEQLGRFALVGEALSVA ATKRLRLLL FAPVACT	660
Qy	661	SLEYNIRVYCLHDTH DALKEVVQ LEKQLGGQLI QEP RVLHF KDSYHNLR LSIH DVPSSLW	720
Db	661	SLEYNIRVYCLHDTH DALKEVVQ LEKQLGGQLI QEP RVLHF KDSYHNLR LSIH DVPSSLW	720
Qy	721	KSKLLVSYQEIPFYHIWNGTQ RY LHCTFTLERVSPSTS DLA CKLWV WQVEGDGQSF SINF	780
Db	721	KSKLLVSYQEIPFYHIWNGTQ QY LHCTFTLERINASTS DLA CKVW WQVEGDGQSF NINF	780
Qy	781	NITKDTRFAELLALESEAGVPA LVGPSAFKIPFLIRQKIISSLDP PCRRGADWRTLAQKL	840
Db	781	NITKDTRFAELLALESEEGGVPA LVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQKL	840
Qy	841	HLD SHLSFFASKP SPTAMI LN LWEARHFPNGNLSQ LAAVAGL GQPDAGLFTV SEAEC	898
Db	841	HLD SHLSFFASKP SPTAMI LN LWEARHFPNGN LGQ LAAVAGL GQPDAGLFTV SEAEC	898

RESULT 5

US-09-970-944-13

; Sequence 13, Application US/09970944

; Publication No. US200

; GENERAL INFORMATION:

; APPLICANT: Herrman, John L

; APPLICANT: Rastelli, Luca

; APPLICANT: Shimkets, Richard A
; TITLE OF INVENTION: No. US20030204052A1 eL Proteins and Nucleic Acids Encoding

Same and

; TITLE OF INVENTION: Anti

; FILE REFERENCE: 21402-138

; CURRENT APPLICATION NUMBER: US/09/970,944

; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/237,862
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-970-944-13

Db 601 EDVLHLGEESPSHLYYCQLEAGACYVFTEQLGRFALVGEALSAATKRLRLLFAPVACT 660
QY 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW 720
Db 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW 720
QY 721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWWQVEGDGQSFSINF 780
Db 721 KSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTSDLACKVWWQVEGDGQSFNINF 780
QY 781 NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPCCRGADWRTLAQKL 840
Db 781 NITKDTRFAELLALESEGGVPALVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQKL 840
QY 841 HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 898
Db 841 HLDSHLSFFASKPSPTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC 898

RESULT 6

US-10-256-702-5

; Sequence 5, Application US/10256702

; Publication No. US20030059859A1

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/256,702
; FILING DATE: 27-Sep-2002
; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/933,261
; FILING DATE: 20-Aug-2001
; APPLICATION NUMBER: 08/808,982
; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 898 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20030059859A1 Relevant
; TOPOLOGY: No. US20030059859A1 Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-256-702-5

Query Match 96.8%; Score 4638; DB 14; Length 898;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 862; Conservative 17; Mismatches 19; Indels 0; Gaps 0;

Qy 1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
Db 1 MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

Qy 61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSQQVEKVFGL 120
Db 61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSGLPTMEVRINVSQQVEKVFGL 120

Qy 121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
Db 121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

Qy 181 VEWLRNEDLVDPSDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
Db 181 VEWLRNEDLVDPSDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSTSAIVY 240

Qy 241 VNGGWSTWTEWSVCASCGRGWQKRSRSCTNPAPLNGAFCEGQNVQKTACATLCPVDGS 300
Db 241 VNGGWSTWTEWSVCASCGRGWQKRSRSCTNPAPLNGAFCEGQNVQKTACATLCPVDGS 300

Qy 301 WSPWSKWSACGLDCTHWRSRECSDPAPRNGGECCQGTLDTRNCSDLCVHSASGPEDVA 360
Db 301 WSSWSKWSACGLDCTHWRSRECSDPAPRNGGECCRGADLDTRNCSDLCCLHTASCPEDVA 360

Qy 361 LYVGLIAVAVCLVLILVLYCRKKEGLDSVDADSSILTSGFQPVSIKPSKADNPILL 420
Db 361 LYIGLVAVAVCLFLLLALGLTYCRKKEGLDSVDADSSILTSGFQPVSIKPSKADNPILL 420

Qy 421 TIQPDLSSTTTTYQGSLCPRQDGSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEEFVS 480
Db 421 TIQPDLSSTTTTYQGSLCSRQDGSPKFQLNGHLLSPLGSGRHTLHHSSPTSEAEDFVS 480

Qy 481 RLSTQNYFRSLPRTGSNMTYGTNFNLGGRLMIPNTGISLLIPPPDAIPRGKIYEIYLTLHK 540
Db 481 RLSTQNYFRSLPRTGSNMYGTNFNLGGRLMIPNTGISLLIPPPDAIPRGKIYEIYLTLHK 540

Qy 541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSW 600
Db 541 PEDVRLPLAGCQTLLSPVVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSW 600

Qy 601 EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSAAKRLKLLLFAFPVACT 660

Db	601 EDVLHLGEESPShLYCQLEAGACYVFTEQLGRFALVGEALSAATKRLRLLLFApVACT 660
Qy	661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW 720
Db	661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW 720
Qy	721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDDLACKLWWQVEGDGQSFSINF 780
Db	721 KSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTSDLACKVWWQVEGDGQSFNINF 780
Qy	781 NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPCCRGADWRTLAQKL 840
Db	781 NITKDTRFAELLALESEGGVPALVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQKL 840
Qy	841 HLDHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 898
Db	841 HLDHLSFFASKPSPTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC 898

RESULT 7

US-10-240-154-16

; Sequence 16, Application US/10240154
; Publication No. US20030175741A1
; GENERAL INFORMATION:
; APPLICANT: Cochran et al.
; TITLE OF INVENTION: SCHIZOPHRENIA RELATED GENES
; FILE REFERENCE: CKFW-P01-006
; CURRENT APPLICATION NUMBER: US/10/240,154
; CURRENT FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: PCT/GB01/01486
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Rattus sp.

US-10-240-154-16

Query Match 96.8%; Score 4638; DB 14; Length 898;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 862; Conservative 17; Mismatches 19; Indels 0; Gaps 0;

Qy	1 MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
Db	1 MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
Qy	61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVFGLE 120
Db	61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSGLPTMEVRINVSRRQQVEKVFGLE 120
Qy	121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
Db	121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
Qy	181 VEWLRNEDLVDPStDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVY 240

Db 181 VEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSTSAAVIVY 240
Qy 241 VNNGWSTWTEWSVCASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGS 300
Db 241 VNNGWSTWTEWSVCASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGS 300
Qy 301 WSPWSKWSACGLDCTHWRSRECSDPAPRNGGECCQGTDLDTRNCTS DLCVHSASGPEDVA 360
Db 301 WSSWSKWSACGLDCTHWRSRECSDPAPRNGGECCRGAIDLDRNCTS DLCCLHTASCPEDVA 360
Qy 361 LYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSVADSSILTSGFQPVSIKPSKADNPHE 420
Db 361 LYIGLVAVAVCLFLLLLALGLIYCRKKEGLDSVADSSILTSGFQPVSIKPSKADNPHE 420
Qy 421 TIQPDLSTTTTYQGSLCPRQDGSPPKFQLTNGHLLSPGGGRHTLHHSSPTSEAEEFVS 480
Db 421 TIQPDLSTTTTYQGSLCSRQDGSPKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEDFVS 480
Qy 481 RLSTQNYFRSLPRGTSNMTRYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHK 540
Db 481 RLSTQNYFRSLPRGTSNMAYGTFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHK 540
Qy 541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCEGSW 600
Db 541 PEDVRLPLAGCQTLLSPVVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCEGSW 600
Qy 601 EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLL FAPVACT 660
Db 601 EDVLHLGEEPSHLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLL FAPVACT 660
Qy 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW 720
Db 661 SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLW 720
Qy 721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSSDLACKLWWVWQVEGDGQSFSINF 780
Db 721 KSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTSDLACKVWWVWQVEGDGQSFSINF 780
Qy 781 NITKDTRFAELLALESEAGVPAVGPSAFKIPFLIRQKIISSLDPCCRGA DWRTLAQKL 840
Db 781 NITKDTRFAELLALESEEGGVPAVGPSAFKIPFLIRQKIIASLDPPCSRGA DWRTLAQKL 840
Qy 841 HLDSHLSFFASKPSPTAMI LNLWEARHF PNGNLSQLAAVAGL GQPDAGLFTVSEAEC 898
Db 841 HLDSHLSFFASKPSPTAMI LNLWEARHF PNGNLSQLAAVAGL GQPDAGLFTVSEAEC 898

RESULT 8

US-10-311-623-1

; Sequence 1, Application US/10311623

; Publication No. US20040023244A1

; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.; GRIFFIN, Jennifer A.

APPLICANT: KALLICK, Deborah A.; TRIBOULEY, Catherine M.

; APPLICANT: YUE, Henry; NGUYEN, Dannie B.

; APPLICANT: TANG, Y. Tom; LAL, Preeti G.

; APPLICANT: POLICKY, Jennifer L.; AZIMZAI, Yalda
 ; APPLICANT: LU, Dyung Aina M.; GRAUL, Richard C.
 ; APPLICANT: YAO, Monique G.; BURFORD, Neil
 ; APPLICANT: HAFALIA, April J. A.; BAUGHN, Mariah R.
 ; APPLICANT: BANDMAN, Olga; ARVIZU, Chandra S.
 ; APPLICANT: YANG, Junming; XU, Yuming
 ; APPLICANT: GANDHI, Ameena R.; WARREN, Bridget A.
 ; APPLICANT: DING, Li; SANJANWALA, Madhusudan M.
 ; APPLICANT: DUGGAN, Brendan M.; LU, Yan
 ; TITLE OF INVENTION: RECEPTORS
 ; FILE REFERENCE: PF-0793 USN
 ; CURRENT APPLICATION NUMBER: US/10/311,623
 ; CURRENT FILING DATE: 2002-12-17
 ; PRIOR APPLICATION NUMBER: US 01/19942
 ; PRIOR FILING DATE: 2001-06-21
 ; PRIOR APPLICATION NUMBER: US 60/214,027
 ; PRIOR FILING DATE: 2000-06-21
 ; PRIOR APPLICATION NUMBER: US 60/228,045
 ; PRIOR FILING DATE: 2000-08-25
 ; PRIOR APPLICATION NUMBER: US 60/255,104
 ; PRIOR FILING DATE: 2000-12-12
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 1
 ; LENGTH: 842
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: Incyte ID No.: US20040023244A1 6052371CD1
 US-10-311-623-1

Query Match 92.1%; Score 4413; DB 16; Length 842;
 Best Local Similarity 93.5%; Pred. No. 0;
 Matches 840; Conservative 1; Mismatches 1; Indels 56; Gaps 1;

Qy	1 MAVRPGLWPALLGIVLAALRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP	60
Db	1 MAVRPGLWPALLGIVLAALRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP	60
Qy	61 VLLVCKAVPATQIFFKCNGEWRQVDHVIERSTDGSSGLPTMEVRINVSRQQVEKVFGL	120
Db	61 VLLVCKAVPATQIFFKCNGEWRQVDHVIERSTDGSSGLPTMEVRINVSRQQVEKVFGL	120
Qy	121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE	180
Db	121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE	180
Qy	181 VEWLRNEDLVDPSDLPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVIVY	240
Db	181 VEWLRNEDLVDPSDLPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVIVY	240
Qy	241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGS	300
Db	241 -----	244
	-----VDGS	
Qy	301 WSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTLDTRNCTSDLCVHSASGPEDVA	360

Db	245	WSPWSKWSACGLDCTHWRRECSDPAPRNGGEECQGTLDTRNCTSDLCVHTASGPEDVA	304
Qy	361	LYVGLIAAVAVCLVLLLVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPILL	420
Db	305	LYVGLIAAVAVCLVLLLVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPILL	364
Qy	421	TIQPDLSTTTTYQGSLCPRQDGSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFFVS	480
Db	365	TIQPDLSTTTTYQGSLCPRQDGSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFFVS	424
Qy	481	RLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIIPPDAIPRGKIYEIYLTLHK	540
Db	425	RLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIIPPDAIPRGKIYEIYLTLHK	484
Qy	541	PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPADSWSLRLKKQSCEGSW	600
Db	485	PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPADSWSLRLKKQSCEGSW	544
Qy	601	EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLLFPVACT	660
Db	545	EDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLLFPVACT	604
Qy	661	SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSLW	720
Db	605	SLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSLW	664
Qy	721	KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDDLACKLWVWQVEGDQFSINF	780
Db	665	KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDDLACKLWVWQVEGDQFSINF	724
Qy	781	NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPCCRGADWRTLAQKL	840
Db	725	NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPCCRGADWRTLAQKL	784
Qy	841	HLDHSLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC	898
Db	785	HLDHSLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC	842

RESULT 9

US-09-970-944-14

; Sequence 14, Application US/09970944

; Publication No. US20030204052A1

; GENERAL INFORMATION:

; APPLICANT: Herrman, John L

; APPLICANT: Rastelli, Luca

; APPLICANT: Shimkets, Richard A

; TITLE OF INVENTION: No. US20030204052A1 Proteins and Nucleic Acids Encoding Same and

; TITLE OF INVENTION: Antibodies Directed Against these Proteins

; FILE REFERENCE: 21402-138

; CURRENT APPLICATION NUMBER: US/09/970,944

; CURRENT FILING DATE: 2002-05-02

; PRIOR APPLICATION NUMBER: 60/237,862

; PRIOR FILING DATE: 2000-10-04

; NUMBER OF SEQ ID NOS: 62

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-944-14

Query Match 59.4%; Score 2845; DB 11; Length 544;
Best Local Similarity 100.0%; Pred. No. 1.5e-236;
Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	358	DVALYVGLIAAVAVCLVLLLLVLILVYCRKKEGLDSVDADSSILTSGFQPVS IKPSKADNP	417
Db	4	DVALYVGLIAAVAVCLVLLLLVLILVYCRKKEGLDSVDADSSILTSGFQPVS IKPSKADNP	63
Qy	418	HLLTIQPDLS TTTTYQGSLCPRQDGSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEE	477
Db	64	HLLTIQPDLS TTTTYQGSLCPRQDGSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEE	123
Qy	478	FVSRSLSTQNYFRSLPRGTSNMTYGTNFNLGGRLMIPNTGISLLI PDAIPRGK IYEIYLT	537
Db	124	FVSRSLSTQNYFRSLPRGTSNMTYGTNFNLGGRLMIPNTGISLLI PDAIPRGK IYEIYLT	183
Qy	538	LHKPEDVRLPLAGCQTLLS PIVSCGPPGVLLTRPVILAMDHCGEPS PDSWSLRLKKQSCE	597
Db	184	LHKPEDVRLPLAGCQTLLS PIVSCGPPGVLLTRPVILAMDHCGEPS PDSWSLRLKKQSCE	243
Qy	598	G SWEDVLHLGEEAPSHLYYCQLEASAC YVFT EQLGRFALVGEALS VAAKRLKLLFAPV	657
Db	244	G SWEDVLHLGEEAPSHLYYCQLEASAC YVFT EQLGRFALVGEALS VAAKRLKLLFAPV	303
Qy	658	ACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHF KDSYHNLRLSIHDVPS	717
Db	304	ACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHF KDSYHNLRLSIHDVPS	363
Qy	718	SLWKS KL LVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS DLACKLWVWQVEGDGQSFS	777
Db	364	SLWKS KL LVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS DLACKLWVWQVEGDGQSFS	423
Qy	778	INFNITKDTRFAELLA LESEAGVPA LVGPSAFKIPFLIRQKIISLDPPCRRGADWRTLA	837
Db	424	INFNITKDTRFAELLA LESEAGVPA LVGPSAFKIPFLIRQKIISLDPPCRRGADWRTLA	483
Qy	838	QKLHLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAGLG QPDAGLFTVSEAE	897
Db	484	QKLHLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAGLG QPDAGLFTVSEAE	543
Qy	898	C 898	
Db	544	C 544	

RESULT 10
US-09-933-261-6
; Sequence 6, Application US/09933261
; Publication No. US20030040046A1
; GENERAL INFORMATION:

APPLICANT: Tessier-Lavigne, Marc
Leonardo, E. David
Hink, Lindsay
Masu, Masayuki
Kazuko, Keino-Masu

TITLE OF INVENTION: Netrin Receptors

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/933,261
FILING DATE: 20-Aug-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/808,982
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC96-217

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 557 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20030040046A1 Relevant
TOPOLOGY: No. US20030040046A1 Relevant

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-933-261-6

Query Match 58.8%; Score 2815.5; DB 10; Length 557;
Best Local Similarity 96.8%; Pred. No. 5.7e-234;
Matches 539; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

Qy 343 NCTSDLCVHSASGPEDVALYVGLIAAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTS 402
Db ||||||| ||:||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 1 NCTSDLXVHTASGPEDVALYVGLIAAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTS 60
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 403 GFQPVSIIKPSKADNPPLLTIQPDLS|||||TYQGSLCPRQDGSPKFQLTNGHLLSPLGGG 462
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 61 GFQPVSIIKPSKADNPPLLTIQPDLS|||||TYQGSLCPRQDGSPKFQLTNGHLLSPLGGG 120
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 463 RHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTNFNLGGRLMIPNTGISLLIP 522
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 121 RHTLHHSSPTSEAEFFSRLSTQNYFRSLPRGTSNMTYGTNFNLGGRLMIPNTGISLLIP 180
Qy 523 PDAIPRGKIYIEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEP 582
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 PDAIPRGKIYIEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEP 240
Qy 583 SPDSWSLRKQSCESWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALS 642
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 SPDSWSLALKQSCESWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALS 300
Qy 643 VAAAKRLKLLLFA PVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFK 702
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 VAAAKRLKLLLFA PVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHLX 360
Qy 703 DSYHNLRLSIHDVPSSLWKSKLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS DLAC 762
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 DSYHNLXLSXHDVPSSLWKSKLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS DLAC 420
Qy 763 KLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISS 822
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 421 KLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISS 480
Qy 823 LDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMI LNLWEARHFPNGNLSQLAAVAG 882
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 481 LDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMI LNLWEARHFPNGNLSQLAAVAG 540
Qy 883 LGQPDAGLFT-VSEAEC 898
| : | | | |
Db 541 TXPAGRWLLSQCSEAEC 557

RESULT 11

US-10-256-702-6

; Sequence 6, Application US/10256702

; Publication No. US20030059859A1

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/256,702

; FILING DATE: 27-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/933,261
; FILING DATE: 20-Aug-2001
; APPLICATION NUMBER: 08/808,982
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20030059859A1 Relevant
; TOPOLOGY: No. US20030059859A1 Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-256-702-6

Query Match 58.8%; Score 2815.5; DB 14; Length 557;
Best Local Similarity 96.8%; Pred. No. 5.7e-234;
Matches 539; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

Qy	343 NCTSDLCVHSASGPEDVALYVGLIAAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTS	402
Db	1 NCTSDLXVHTASGPEDVALYVGLIAAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTS	60
Qy	403 GFQPVSIKPSKADNPHELTIQPDLS TTTYQGSCLPRQDGSPKFQLTNGHLLSPLGGG	462
Db	61 GFQPVSIKPSKADNPHELTIQPDLS TTTYQGSCLPRQDGSPKFQLTNGHLLSPLGGG	120
Qy	463 RHTLHHSSPTSEAEFFSRLSTQNYFRSLPRGTSNMTYGTNFLLGRLMPNTGISLLIP	522
Db	121 RHTLHHSSPTSEAEFFSRLSTQNYFRSLPRGTSNMTYGTNFLLGRLMPNTGISLLIP	180
Qy	523 PDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPEP	582
Db	181 PDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPEP	240
Qy	583 SPDSWSLRKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALS	642
Db	241 SPDSWSLALKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALS	300
Qy	643 VAAAKRLKLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFK	702
Db	301 VAAAKRLKLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHLX	360
Qy	703 DSYHNLRLSIHDVPSSLWKSKLIVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLAC	762
Db	361 DSYHNLXLSXHDVPSSLWKSKLIVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLAC	420
Qy	763 KLWWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISS	822

Db ||||||| 421 KLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISS 480
Qy 823 LDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAG 882
Db ||||||| 481 LDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAG 540
Qy 883 LGQPDAGLFT-VSEAEC 898
| : |||||
Db 541 TXPAGRWLLSQCSEAEC 557

RESULT 12

US-09-970-944-15

; Sequence 15, Application US/09970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
; APPLICANT: Herrman, John L
; APPLICANT: Rastelli, Luca
; APPLICANT: Shimkets, Richard A
; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same and
; TITLE OF INVENTION: Antibodies Directed Against these Proteins
; FILE REFERENCE: 21402-138
; CURRENT APPLICATION NUMBER: US/09/970,944
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/237,862
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 931
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans

US-09-970-944-15

Query Match 58.2%; Score 2787; DB 11; Length 931;
Best Local Similarity 57.3%; Pred. No. 3.6e-231;
Matches 522; Conservative 153; Mismatches 208; Indels 28; Gaps 9;

Qy 9 PALLGIVLAAWLRGSGAQQA---TVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLVLC 65
| | :|:| || || : | |: |||||:|||: ||||||| | |
Db 26 PAL--ALLSASGTGSAAQDDEFFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYC 83

Qy 66 KAVPATQIFFKCNGEWWRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVFGLEEYWCQ 125
| | |||||:||||| ||| | ||||: | :||| | || | :|||||:||| |:|||
Db 84 KASPAHQIYFKCNSEWHQKDHVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQ 143

Qy 126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPPAEVEWLR 185
| |||||:|||:||| | ||| | ||||| | ||||| ::| | ||||| | |||||:
Db 144 CVAWSSAGTTKSQKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK 203

Qy 186 NEDLVDPSDLPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVIVVNGGW 245
| ||::|: | | ||| :|:|:|:|:|:|:|:|:|:|:|:|:| | ||||| | |
Db 204 NEDIIDPAEDRNFYITIDHNLIKQARLSDTANYTCVAKNIVAKRKSTTATVIVVNGGW 263

Qy 246 STWTEWSVCSASCGRGWQKRSRSTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWPWS 305

Db	264	STWTEWSVCNSRCGRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIAC TTLCPVDGRWTWS 323
Qy	306	KWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTS DLCVHSASGPEDVALYVGL 365
Db	324	KWSTCGTECTHWRRRECTAPAPKNGKDCDGLVLQSKNCTDGLCMQAAPDSDDVALYVGI 383
Qy	366	-IAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPHLLTIQP 424
Db	384	VIAVTVC LAITVVVALFVYRKHNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVPP 440
Qy	425	DLSTTTTYQGSCLCPRQDGSPPKFQLTNGHLLSPLGGGRHTLHHSS---PTSEAEFVS 480
Db	441	DITSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFSS 499
Qy	481	RLS---TQNYF-----RSLPRGT--SNMTYGTTFNFLGGRLMIPNTGISLLIPPDAI 526
Db	500	KLSPQMTQSLLNEALNLKNQSLARQTDPSC TAFGT FNSLGGHLIIPNSGVSLI PAGAI 559
Qy	527	PRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHC GE PSPDS 586
Db	560	PQGRVYEMYVTVHRKENMRPPMEDSQ TLLTPV VSCGPPGALLTRPVILTLHHCADPSTED 619
Qy	587	WSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA 646
Db	620	WKIQLKNQAVQGQWEDVVVGEENFTTPCYIQ LDAEACHILTENLSTYALVGQSTTKAAA 679
Qy	647	KRLKLLLFA PVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYH 706
Db	680	KRLKLAIFGPLCCSSLEYSIRVYCLDDTQ DALKEV LQLERQMGGQ LLEEPKALHF KGSIH 739
Qy	707	NLRLSIHDVPSLWKS KLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS D LACKLWV 766
Db	740	NLRLSIHDIAHSLWKS KLLAKYQEIPFYHIWSGSQRNLHCTFTLERLSLNTVELVCKLCV 799
Qy	767	WQVEGDGQSFSINFNITKDTRFAELLA SEAGVPA LVGPSAFKIPFLIRQKIISSL DPP 826
Db	800	RQVEGEQI FQLNCTVSEEPTGIDLP LLDPASTITTVTGPSAFSIP LPIRQKLCSSLDAP 859
Qy	827	CRRGADWRTL A QKLHLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQ LAAVAGLGQP 886
Db	860	QTRGHDRMLAHKLNLD RYLNYFATKSSPTGVILD LWEAQNF PDGNLSMLA AVLEEMGRH 919
Qy	887	DAGLFTVSEAE 897
Db	920	ETVVSLAAEGQ 930

RESULT 13

US-10-087-684-35

; Sequence 35, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: MacDougall, John R.
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen

; APPLICANT: Stone, David J.
; APPLICANT: Grosse, William M.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Cathereine E.
; APPLICANT: Casman, Stacie, J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Guo, Xiaojaia
; APPLICANT: Miller, Charles E.
; APPLICANT: Gangolli, Esha A.
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-214 CIP
; CURRENT APPLICATION NUMBER: US/10/087,684
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 35
; LENGTH: 931
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-087-684-35

Query Match 58.2%; Score 2787; DB 12; Length 931;
Best Local Similarity 57.3%; Pred. No. 3.6e-231;
Matches 522; Conservative 153; Mismatches 208; Indels 28; Gaps 9;

Qy 9 PALLGIVLAAWLRGSGAQQSA---TVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLVC 65
||| :||| |||| : | | |:||||:|||:||||||| |||
Db 26 PAL--ALLSASGTGSAAQDDEFFHELPETFPSPPEPLPHFLIEPEEAYIVKNKPVNLYC 83

Qy 66 KAVPATQIFFKCNGEWRQVDHIERSTDGSSGLPTMEVRINVSRRQQVEKVGLEEWQC 125
|| |||||:|||| ||| | |||: | :||| ||| :|||||:||| :|||
Db 84 KASPAHQIYFKCNSEWVHQKDHWVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQ 143

Qy 126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLR 185

Db	144	CAVASSAGTTSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK	203
Qy	186	NEDLVDPSDLPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAASAVIVYVNNGW	245
Db	204	NEDIIDPAEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNNGW	263
Qy	246	STWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGWSPW	305
Db	264	STWTEWSVCNSRCGRGYQKRRTCTNPAPLNGGAFCEGQSVQKIACTTLCPVDGRTSWS	323
Qy	306	KWSACGLDCTHWRSRECSDPAPRNGGECCQGTDLDRNCTS DL CVHSASGPEDVALYVGL	365
Db	324	KWSTCGTECTHWRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPDSDDVALYVG	383
Qy	366	-IAVAVCLVLLLLVLILVYCRKKEGLDSVDSSILTSGFQPVSIKPSKADNPHLLTIQP	424
Db	384	VIAVTVCLAITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVPP	440
Qy	425	DLSTTTTYQGSCLPRQDGSPKFQLTNHGILLSPLGGGRHTLHHSS---PTSEAEFVS	480
Db	441	DLTSAAMAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIVYNSSGAVTPQDDLAEFSS	499
Qy	481	RLS---TQNYF-----RSLPRGT--SNMTYGTNFLLGGRLMIPNTGISLLIPPDAI	526
Db	500	KLSPQMTQSLLNEALNLKNQSLARQTDPSCATAFGTFNLSGGHLIIPNSGVSLIPAGAI	559
Qy	527	PRGIYEIYLTLHKPEDVRLPLAGCQTLISPIVSCCGPPGVLLTRPVILAMDHCGEPS	586
Db	560	PQGRVYEMYVTVHRKENMRPPMEDSQTLTLPVVSCGPPGALLTRPVILTLHHCADPSTE	619
Qy	587	WSLRKKQSCEGSWEDVHLGEEAPS HLYYCQLEASACYVFTEQLGRFALVGEALSVAAA	646
Db	620	WKIQLKNOAVQGQWEDVVVVGEENFTTPCYIQDAEACHILTENLSTYALVGQSTTKAAA	679
Qy	647	KRLKLLLFA PVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRLHFKD SYH	706
Db	680	KRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEEPKALHFKGSIH	739
Qy	707	NLRLSIHDVPSLWKS KLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS DLACKLWV	766
Db	740	NLRLSIHDIAHSIWKSLAKYQEIPFYHIWGSQRNLHCTFTLERLSINTVELVKLCV	799
Qy	767	WQVEGDGQSFSINFNITKDT RFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPP	826
Db	800	RQVEGEGQIFQLNCTVSEEPTGIDPLLLDPASTITTVGPSAFSIPLPIRQKLCSSLDAP	859
Qy	827	CRRGADWRTL AOKLHLD SHLSFFASKPSPTAMILNLWEARHF PNGNLSQLAAVAGLGQP	886
Db	860	QTRGHDWMLAHKLNLDRYLNYFATKSSPTGVILDWEAQNFPDGNLSMLAAVLEEMGRH	919
Qy	887	DAGLFTVSEAE 897	
Db	920	ETVVSLAEGQ 930	

US-09-972-211-121
; Sequence 121, Application US/09972211
; Publication No. US20040048245A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Spytek, Kimberly Ann
; APPLICANT: Li, Li
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John R
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David J
; APPLICANT: Smithson, Glennda
; APPLICANT: Szekeres Jr, Edward S
; TITLE OF INVENTION: No. US20040048245A1 Human Proteins, Polynucleotides
Encoding Them And
; TITLE OF INVENTION: Methods Of Using The Same
; FILE REFERENCE: 21402-141
; CURRENT APPLICATION NUMBER: US/09/972,211
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/238,325
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,323
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/238,400
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,397
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,401
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,379
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,402
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 30/238,384
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,373
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,372
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,383
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,382
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/275,892
; PRIOR FILING DATE: 2001-03-14

; PRIOR APPLICATION NUMBER: 60/296,860
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 198
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 121
; LENGTH: 931
; TYPE: PRT
; ORGANISM: *Mus musculus*
US-09-972-211-121

Query Match 58.2%; Score 2787; DB 12; Length 931;
 Best Local Similarity 57.3%; Pred. No. 3.6e-231;
 Matches 522; Conservative 153; Mismatches 208; Indels 28; Gaps 9;
 Qy 9 PALLGIVLAAWLRGSGAQQA---TVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC 65
 ||| :|:| ||| | : | |: ||||:|||: ||||||| |||
 Db 26 PAL--ALLSASGTGSAAQDDEFFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYC 83
 Qy 66 KAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVGLEEWYCQ 125
 || |||||:||| | ||| | |||:: | :||| ||| :|||:|||:||| | :|||
 Db 84 KASPATQIYFKCNSEWHQKDHHVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQ 143
 Qy 126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLR 185
 ||| :|||:|||:|||:||| ||| ||| ||| ||| :||| ||| :||| ||| :|||:
 Db 144 CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVVAEVEWLK 203
 Qy 186 NEDLVDPSDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVIVYVNGGW 245
 |||:|||: | | ||| :|:|||:|||:|||:|||:|||:|||:|||:||| :| |||:|||
 Db 204 NEDIIDPAEDRNFYITIDHNLIQARLSDTANYTCVAKNIVAKRKTATVIVIVYVNGGW 263
 Qy 246 STWTEWSVCSASCGRGWQKRSRSTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWS 305
 |||:|||: | |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 264 STWTEWSVCNSRCGRGYQKRRTCTNPAPLNGGAFCEGQSVQKIACTLCPVDGRWTWS 323
 Qy 306 KWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTS DLCVHSASGPEDVALYVGL 365
 ||| ||| :||| | |||: |||:|||:|||:||| | | :||| |||:||| :| :|||:|||:
 Db 324 KWSTCGTECTHWRRRECTAPAPKNGGKDCDGLVLQSKNTDGLCMQAAPDSDDVALYVGI 383
 Qy 366 -IAAVAVCLVLLLLVLILVYCRKKEGLSDVADSSIITSGFQPVSIKPSKADNPHTLIQP 424
 ||| ||| : :| : ||| : | :||| | |||:||| :| | ||| :|
 Db 384 VIAVTVCLAITVVVALFVYRKNRDFESDIIDSSALNGGFQPVNIKAARQD--LLAVPP 440
 Qy 425 DLSTTTTYQGSLCPRQDGSPPKFQLTNGHLLSPGGGRHTLHHSS----PTSEAEFVS 480
 |||: | :| : | | | :||| :||| :| :||| | :| :|||
 Db 441 DLTSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFSS 499
 Qy 481 RLS---TQNYF-----RSLPRGT---SNMTYGTNFLLGGRLMIPNTGISLLIPPPDAI 526
 :||| | :| :| | | :| | | :| | | | | :| :| :| :| :|
 Db 500 KLSPQMTQSLLNEALNLKNQSLARQTDPSCATAFGTFNSLGHHLIIPNSGVSLIPAGAI 559
 Qy 527 PRGKIYEIYLTLHKPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPSD 586
 | :|:|||:|||:|||: | :| | | | :| | | | | | :| :| :| :|
 Db 560 PQGRVYEMYVTVHRKENMRPPMEDSQTLTPVSCGPPGALLTRPVILTLHHCADPSTED 619
 Qy 587 WSLRLKKQSCEGSWEDVLHLGEEAPSHLYCQLEASACYVFTEQLGRFALVGEALSVAAA 646
 | :| | | :| | | | :| | | | | | | | | | :| :| :| :| :|
 Db 620 WKIQOLKNOAVOGOWEDVVVVGEENFTTPCYIOLDAAEACHILTENLSTYALVGOSTTKAAA 679

Qy	647	KRLKLLLFA PVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYH	706
	: : : : : : :		
Db	680	KRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQMGQLLEEPKALHFKGSIH	739
Qy	707	NLRLSTHDV PSSIWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS DLACKLWV	766
	: : : :		
Db	740	NLRLSIHDIAHSLWKS KLAKYQEIPFYHIWSGSQRNLHCTFTLERLS LNTVELVCKLCV	799
Qy	767	WQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKISSL DPP	826
	: : :::: : : : : :		
Db	800	RQVEGEQI FQLNCTVSEEPTGIDLP LLDPASTITTVTGPSAFS IPLPIRQKLCSLDAP	859
Qy	827	CRRGADWR TLAQKLHLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAA VAGLGQP	886
	: : : : : : : : : : :		
Db	860	QTRGH DW RMLAHKLNLDRYLN YFATKSSPTGVILD LWEAQNF PDGNLSMLAA VLEEMGRH	919
Qy	887	DAGLFTVSEAE	897
	: : : :		
Db	920	ETVVSLAAEGQ	930

RESULT 15

US-10-037-417-117

; Sequence 117, Application US/10037417

; Publication No. US20040052806A1

; GENERAL INFORMATION:

; APPLICANT: Kekuda, Ramesh
; APPLICANT: Alsobrook II, John P
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Patturajan, Meera
; APPLICANT: Grosse, William M
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Li, Li
; APPLICANT: Gorman, Linda
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Sciore, Paul
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Guo, Xiao jia
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Anderson, David W
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Miller, Charles E
; APPLICANT: Eisen, Andrew J
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-235
; CURRENT APPLICATION NUMBER: US/10/037,417
; CURRENT FILING DATE: 2002-09-20

; PRIOR APPLICATION NUMBER: 60/260,018
 ; PRIOR FILING DATE: 2001-01-05
 ; PRIOR APPLICATION NUMBER: 60/260,360
 ; PRIOR FILING DATE: 2001-01-08
 ; PRIOR APPLICATION NUMBER: 60/272,411
 ; PRIOR FILING DATE: 2001-02-28
 ; PRIOR APPLICATION NUMBER: 60/272,817
 ; PRIOR FILING DATE: 2001-03-02
 ; PRIOR APPLICATION NUMBER: 60/291,186
 ; PRIOR FILING DATE: 2001-05-15
 ; PRIOR APPLICATION NUMBER: 60/303,231
 ; PRIOR FILING DATE: 2001-07-05
 ; PRIOR APPLICATION NUMBER: 60/305,060
 ; PRIOR FILING DATE: 2001-07-12
 ; PRIOR APPLICATION NUMBER: 60/318,405
 ; PRIOR FILING DATE: 2001-09-10
 ; PRIOR APPLICATION NUMBER: 60/318,700
 ; PRIOR FILING DATE: 2001-09-12
 ; NUMBER OF SEQ ID NOS: 227
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 117
 ; LENGTH: 931
 ; TYPE: PRT
 ; ORGANISM: *Caenorhabditis elegans*
 US-10-037-417-117

Query Match 58.2%; Score 2787; DB 12; Length 931;
 Best Local Similarity 57.3%; Pred. No. 3.6e-231;
 Matches 522; Conservative 153; Mismatches 208; Indels 28; Gaps 9;

QY	9 PALLGIVLAAWLRGSGAQQSAA--TVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLVC 65
	: : : : : :
Db	26 PAL--ALLSASGTGSAAQDDEFFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYC 83
QY	66 KAVPATQIFFKCNGEWRQVDHVIERSTDGSSGLPTMEVRINVSQQVEKVFGLEEYWCQ 125
	: : : : : :
Db	84 KASPAHQIYFKCNSEWHQKDHVVDERVDETSGLIVREVSIEISRQVEELFGPEDYWCQ 143
QY	126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWL 185
	: : : :: :
Db	144 CVAWSSAGTTKSQKAYVRIAYLRKTQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK 203
QY	186 NEDLVDPSDLPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVIVYVNGGW 245
	:: : : : : : : : : : : : : :
Db	204 NEDIIDPAEDRNFYITIDHNLIICKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW 263
QY	246 STWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWS 305
	: : : : : : : : : : : : : :
Db	264 STWTEWSVCNSRCGRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIACTTLCPWDGRWTSWS 323
QY	306 KWSACGLDCTHWRSRECSDPAPRNGGEECQGTLDTRNCTS DLCVHSASGPEDVALYVGL 365
	: : : : : : : : : :
Db	324 KWSTCGTECTHWRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPDSDDVALYVGI 383
QY	366 -IAVAVCLVLLLVILVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPHLTIQP 424
	: : : : : : :
Db	384 VIAVTCLAITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVPP 440

Qy	425	DLSTTTTYQGSLCPRQDGSPKFQLTNGHLLSPLGGGRHTLHHSS---PTSEAEFVS	480
	:: : : : : : : : :		
Db	441	DLTSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNSSGAVT	PQDDLAEFSS 499
Qy	481	RLS---TQNYF-----RSLPRGT--SNMTYGTNFLGGRLMPINTGISLLIPDAI	526
	: : : : : :		
Db	500	KLSPQMTQSLLNEALNLKNQSLARQTDPSCATAFGTFNSLGGHLIIPNSGVSLIPAGAI	559
Qy	527	PRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHC	GEPSD 586
	: :: : : : : : : : : : : : : : : :		
Db	560	PQGRVYEMYVTVHRKENMRPPMEDSQTLTPVVSCGPPGALLTRPVILTLHHCADP	STED 619
Qy	587	WSLRKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALS	VAAA 646
	: : : :		
Db	620	WKIQLKNQAVQGWEDVVVGEENFTTPCYIQLDAEACHILTENLSTYALVGQSTTKAA	A 679
Qy	647	KRLKLLLFAVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEP	RVLHF KDSYH 706
	: :		
Db	680	KRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQMGQLLEEP	KALHF KGSIH 739
Qy	707	NLRLSIHDVPSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSP	STSDLACKLWV 766
	:		
Db	740	NLRLSIHDIASHLWKSLLAKYQEIPFYHIWSGSQRNLLHCTFTLERLS	LNTVELVCKLCV 799
Qy	767	WQVEGDGQSFISINFNIKTDRFAELLALESAGVPALVGPSAFKIPFLIRQ	KIISSL DPP 826
	:		
Db	800	RQVEGEGQIFQLNCTVSEEPTGIDLPLLD PASTITTVGPSAFSIPLPIRQ	KLCSSLDAP 859
Qy	827	CRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEARHF	PNGNLSQLAAVAGLGQP 886
Db	860	QTRGHDWMLAHKLNLDRYLN YFATKSSPTGVILD LWEAQNF	PDGNLSMLAAVLEEMGRH 919
Qy	887	DAGLFTVSEAE	897
	: :		
Db	920	ETVVSLAEGQ	930

Search completed: July 6, 2004, 14:37:22
 Job time : 58 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2004, 14:29:11 ; Search time 49 Seconds
(without alignments)
5782.354 Million cell updates/sec

Title: US-10-624-932-2

Perfect score: 4791

Sequence: 1 MAVRPGLWPALLGIVLAAWL.....AVAGLGQPDAGLFTVSEAE 898

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:
1: sp_archea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phage:
10: sp_plant:
11: sp_rat:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriap:
17: sp_archeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description

1	4685	97.8	898	11	Q8K1S4	Q8kls4 mus musculu
2	4638	96.8	898	11	O08721	O08721 rattus norv
3	2845	59.4	544	4	Q96GP4	Q96gp4 homo sapien
4	2787	58.2	931	11	O08747	O08747 mus musculu
5	2767.5	57.8	950	11	Q8CD16	Q8cd16 mus musculu
6	2761	57.6	931	13	Q7T2Z5	Q7t2z5 gallus gall
7	2755	57.5	931	4	O95185	O95185 homo sapien
8	2646.5	55.2	943	13	Q8JGT4	Q8jgt4 xenopus lae
9	2585	54.0	1008	11	Q80Y85	Q80y85 mus musculu
10	2578.5	53.8	945	11	Q8K1S3	Q8kls3 mus musculu
11	2578.5	53.8	945	11	O08722	O08722 rattus norv
12	2572.5	53.7	945	11	Q9D398	Q9d398 mus musculu
13	2566	53.6	934	4	Q8IZJ1	Q8izj1 homo sapien
14	2558.5	53.4	945	4	Q86SN3	Q86sn3 homo sapien
15	2200	45.9	956	11	Q8K1S2	Q8kls2 mus musculu
16	2189.5	45.7	948	4	Q8WYP7	Q8wyp7 homo sapien
17	1668.5	34.8	597	4	Q8IUT0	Q8iut0 homo sapien
18	1458	30.4	328	11	Q80T71	Q80t71 mus musculu
19	1242.5	25.9	554	4	Q8N1Y2	Q8n1y2 homo sapien
20	997	20.8	1072	5	Q9NBL0	Q9nb10 drosophila
21	992	20.7	1072	5	Q9V7B5	Q9v7b5 drosophila
22	981.5	20.5	366	4	Q9H9F3	Q9h9f3 homo sapien
23	980	20.5	947	5	Q26262	Q26262 caenorhabdi
24	977	20.4	947	5	O44171	O44171 caenorhabdi
25	692	14.4	199	13	Q9PVD5	Q9pvd5 petromyzon
26	552.5	11.5	351	4	Q8TF26	Q8tf26 homo sapien
27	377.5	7.9	2673	4	Q96SC3	Q96sc3 homo sapien
28	377.5	7.9	5636	4	Q96RW7	Q96rw7 homo sapien
29	318	6.6	325	5	Q8I1K1	Q8i1k1 drosophila
30	300	6.3	518	4	Q8IV45	Q8iv45 homo sapien
31	293	6.1	1172	11	Q8CG21	Q8cg21 mus musculu
32	293	6.1	1172	11	Q7TMT3	Q7tmt3 mus musculu
33	292	6.1	1582	11	Q8CGM0	Q8cgm0 mus musculu
34	286	6.0	1081	5	Q9U631	Q9u631 drosophila
35	285	5.9	1083	5	Q9VTT0	Q9vtt0 drosophila
36	285	5.9	1091	5	Q7YU67	Q7yu67 drosophila
37	276	5.8	1461	5	Q8MYA8	Q8mya8 caenorhabdi
38	275.5	5.8	1122	11	Q7TT33	Q7tt33 mus musculu
39	275	5.7	1522	11	Q80ZF8	Q80zf8 mus musculu
40	274.5	5.7	1573	4	Q8NGW8	Q8ngw8 homo sapien
41	273.5	5.7	478	11	Q8BVE5	Q8bve5 mus musculu
42	271.5	5.7	685	6	Q9TTS5	Q9tts5 bos taurus
43	271.5	5.7	5146	6	Q8SPM4	Q8spm4 bos taurus
44	271	5.7	1560	11	Q8CGM1	Q8cgm1 mus musculu
45	270.5	5.6	1171	11	Q8CGB2	Q8cgb2 mus musculu

ALIGNMENTS

RESULT 1

Q8K1S4

ID Q8K1S4 PRELIMINARY; PRT; 898 AA.
AC Q8K1S4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Netrin receptor Unc5h1.
 GN UNC5H1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Engelkamp D.;
 RT "Cloning of three mouse unc-5 genes and their expression patterns at
 mid-gestation.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AJ487852; CAD32250.1; -.
 DR MGD; MGI:894682; Unc5h1.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00090; tsp_1; 2.
 DR Pfam; PF00791; ZU5; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00209; TSP1; 2.
 DR SMART; SM00218; ZU5; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR PROSITE; PS50092; TSP1; 2.
 KW Receptor.
 SQ SEQUENCE 898 AA; 98856 MW; 59F04BA2E196C1DB CRC64;

 Query Match 97.8%; Score 4685; DB 11; Length 898;
 Best Local Similarity 96.7%; Pred. No. 0;
 Matches 868; Conservative 19; Mismatches 11; Indels 0; Gaps 0;

 Qy 1 MAVRPGILWPALLGIVLAAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
 |||||||
 Db 1 MAVRPGILWPALLGIVLTAWLRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

 Qy 61 VLLVCKAVPATQIFFKCNGEWRQVDHVIERSTDGSSGLPTMEVRINVSROQVEKVFGL 120
 |||||||
 Db 61 VLLVCKAVPATQIFFKCNGEWRQVDHVIERSTDGSSGLPTMEVRINVSROQVEKVFGL 120

 Qy 121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
 |||||||
 Db 121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

 Qy 181 VEWLRNEDLVDPSDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
 |||||||
 Db 181 VEWLRNEDLVDPSDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240

 Qy 241 VNGGWSTWTEWSVCASCGRGWQKRSRSCTNPAPLNGGAFCEGQNQVKTACATLCPVDGS 300

Db	241	VNGGWSTWTEWSVCSASCGRGWQKRSRSCNPAPLNGAFCEGQNVQKTACATLCPVDGS	300
Qy	301	WSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTLDTRNCTS DL CVHSASGPEDVA	360
Db	301	WSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECRGADLDTRNCTS DL CLHTSSGPEDVA	360
Qy	361	LYVGLIAVAVCLVLLLV LILVYCRKKEGLDS DVADSSILTSGFQPVS IKPSKADNP HLL	420
Db	361	LYIGLVA VAVCL LLLLVL VLVIYCRKKEGLDS DVADSSILTSGFQPVS IKPSKADNP HLL	420
Qy	421	TIQPDLS TTTTYQGS LCPRQDGPSPKFQLTNGHLLSPLGGGRH TLHHSSPTSEAEEF VS	480
Db	421	TIQPDLS TTTTYQGS LCPRQDGPSPKFQLSNGHLLSPLGS GRH TLHHSSPTSEAEDF VS	480
Qy	481	RLSTQNYFRSLP RGT SNMTY GTFNFLGGRLM PNTGISL LIPPDAI PRGKI YEIYL TLHK	540
Db	481	RLSTQNYFRSLP RGT SNMAY GTFNFLGGRLM PNTGISL LIPPDAI PRGKI YEIYL TLHK	540
Qy	541	PEDVRLPLAGC QTLLS PIVSCGPPGVLLTRPV ILAMDHC GEPS PD SWSL RLKKQS CEGSW	600
Db	541	PEDVRLPLAGC QTLLS PIVSCGPPGVLLTRPV ILAMDHC GEPS PD SWSL RLKKQS CEGSW	600
Qy	601	EDVLHLGEEAPSH LYCYQLEASACYVFT EQLGRF ALVGEAL SVAAKRLK LLLFAPVACT	660
Db	601	EDVLHLGEEAPSH LYCYQLEAGACYVFT EQLGRF ALVGEAL SVAATKRLR LLLFAPVACT	660
Qy	661	SLEYNIRVYCLHDTHDALKEVVQ LEKQLGGQLI QEP RVLHF KDSYHN LRLSI HDVPSSI W	720
Db	661	SLEYNIRVYCLHDTHDALKEVVQ LEKQLGGQLI QEP RVLHF KDSYHN LRLSI HDVPSSI W	720
Qy	721	KSKLLVSYQEIPFYHIWNGTQ RY LHCTFTL ERVSP STSDLACKL WVQVEGDGQ SF SINF	780
Db	721	KSKLLVSYQEIPFYHIWNGTQ QY LHCTFTL ERVNASTS DLACKV WWQVEGDGQ SF NINF	780
Qy	781	NITKDTRFAELL ALESEAGVPA LVGPSAFKIPFLIRQKIISSL DPPC RRGADWRTLAQKL	840
Db	781	NITKDTRFAEMLALESEGGVPA LVGPSAFKIPFLIRQKIISSL DPPC SRGADWRTLAQKL	840
Qy	841	HLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAA VAGLGQPDAGLFTVSEAEC	898
Db	841	HLD SHLSFFASKPSPTAMILNLWEARHFPNGNLGQLAAA VAGLGQPDAGLFTVSEAEC	898

RESULT 2

008721

ID 008721 PRELIMINARY; PRT; 898 AA.
 AC 008721;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Transmembrane receptor UNC5H1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE=Brain, and Ventral spinal cord;
 RX MEDLINE=97271897; PubMed=9126742;
 RA Leonardo E.D., Hinck L., Masu M., Keino-Masu K., Ackerman S.L.,
 RA Tessier-Lavigne M.;
 RT "Vertebrate homologues of C. elegans UNC-5 are candidate netrin
 receptors.";
 RL Nature 386:833-838(1997).
 DR EMBL; U87305; AAB57678.1; -.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00090; tsp_1; 2.
 DR Pfam; PF00791; ZU5; 1.
 DR PRINTS; PRO1705; TSP1REPEAT.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00209; TSP1; 2.
 DR SMART; SM00218; ZU5; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR PROSITE; PS50092; TSP1; 2.
 KW Receptor.
 SQ SEQUENCE 898 AA; 98840 MW; 7A3C9E7ACA135 CRC64;

 Query Match 96.8%; Score 4638; DB 11; Length 898;
 Best Local Similarity 96.0%; Pred. No. 0;
 Matches 862; Conservative 17; Mismatches 19; Indels 0; Gaps 0;

 Qy 1 MAVRPGIWPALLGIVLAALRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 1 MAVRPGIWPVLLGIVLAALRGSGAQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP 60

 Qy 61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVFGL 120
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 61 VLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSGLPTMEVRINVSRRQQVEKVFGL 120

 Qy 121 EYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 121 EYWCQCVAWSSSGTTKSQKAYIRIAYLRLRNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE 180

 Qy 181 VEWLRNEDLVDPSDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVIVY 240
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 181 VEWLRNEDLVDPSDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSTSAAIVIVY 240

 Qy 241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGS 300
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGS 300

 Qy 301 WSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTLDTRNCTS DLCVHSASGPEDVA 360
 || ||||||| ||||||| ||||||| ||||||| :| ||||||| ||||| :| :|| |||||

Db 301 WSSWSKWSACGLDCTHWRSRECSDPAPRNGGEECRGADLDTRNCTS DLCLHTASCPEDVA 360
 Qy 361 LYVGLIAVAVCLVLLLVLILVYCRKKEGLDSVADSSILTSGFQPVS IKPSKADNP HLL 420
 ||:||:||||| |||| | |;||||| ||||||| ||||||| ||||||| |||||||
 Db 361 LYIGLVA VAVCLFLLL ALGLIYCRKKEGLDSVADSSILTSGFQPVS IKPSKADNP HLL 420
 Qy 421 TIQPDLSTTTTYQGS LCP RDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEA EEFV S 480
 ||||||| ||||||| ||||| :||| ||||| ||||||| ||||||| :|||
 Db 421 TIQPDLSTTTTYQGS LCSRQDG PSPKFQLSNGHLLSPLGSGRHTLHHSSPTSEA EDFV S 480
 Qy 481 RLSTQNYFRSLPRGTSNM TYGTFNFLGGRLMIPNTGISL I PPDAI PRGKIYEIYLTLHK 540
 ||||||| ||||||| ||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 481 RLSTQNYFRSLPRGTSNM MAYGTFNFLGGRLMIPNTGISL I PPDAI PRGKIYEIYLTLHK 540
 Qy 541 PEDVRLPLAGC QTLLSPIVSCGPGVLLTRPVILAMDHC GEPSPDSW SLRLKKQS CEGSW 600
 ||||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 541 PEDVRLPLAGC QTLLSPVVSCGPGVLLTRPVILAMDHC GEPSPDSW SLRLKKQS CEGSW 600
 Qy 601 EDVLHLGEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLL FAPVACT 660
 ||||||| :||| ||||| ||||| ||||| ||||| ||||| :|||
 Db 601 EDVLHLGEESP SHLYYCQLEAGACYVFTEQLGRFALVGEALSVAATKRLRLLL FAPVACT 660
 Qy 661 SLEYNIRVYCLHDTHDALKEV VQLEKQLGGQLIQEPRVLHF KDSYHNLRLSIHDV PSSLW 720
 ||||||| ||||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 661 SLEYNIRVYCLHDTHDALKEV VQLEKQLGGQLIQEPRVLHF KDSYHNLRLSIHDV PSSLW 720
 Qy 721 KSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS DLA CKLWWQVEGDGQSF SINF 780
 ||||||| :||| ||||| :||| :||| :||| :||| :||| :|||
 Db 721 KSKLLVSYQEIPFYHIWNGTQQYLHCTFTLERINASTS DLA CKVWWQVEGDGQSF NINF 780
 Qy 781 NITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSL DPPCRRGADWRT LAQKL 840
 ||||||| ||||| ||||| ||||| :||| :||| :||| :|||
 Db 781 NITKDTRFAELLALESEGGVPALVGPSAFKIPFLIRQKIIASLDPPCSRGADWRT LAQKL 840
 Qy 841 HLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 898
 ||||||| ||||||| ||||| ||||| ||||| |||||
 Db 841 HLD SHLSFFASKPSPTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC 898

RESULT 3

Q96GP4

ID Q96GP4 PRELIMINARY; PRT; 544 AA.
 AC Q96GP4;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Similar to transmembrane receptor Unc5H1 (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; BC009333; AAH09333.1; -

DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00791; ZU5; 1.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00218; ZU5; 1.
 KW Receptor; Transmembrane.
 FT NON_TER 1 1
 SQ SEQUENCE 544 AA; 59949 MW; 350A7BA53375CCAECRC64;

Query Match 59.4%; Score 2845; DB 4; Length 544;
 Best Local Similarity 100.0%; Pred. No. 1.9e-253;
 Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	358	DVALYVGLIAAVAVCLVLLLLVLILVYCRKKEGLDSVDADSSILTSGFQPVSIKPSKADNP	417
Db	4	DVALYVGLIAAVAVCLVLLLLVLILVYCRKKEGLDSVDADSSILTSGFQPVSIKPSKADNP	63
Qy	418	HLLTIQPDLSTTTTYQGSCLCPRQDGSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAAEE	477
Db	64	HLLTIQPDLSTTTTYQGSCLCPRQDGSPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAAEE	123
Qy	478	FVSRLSTQNYFRSLPRTGSNMTYGTNFNLGGRLMIPNTGISLLIIPPDAIPRGKIYEIYLT	537
Db	124	FVSRLSTQNYFRSLPRTGSNMTYGTNFNLGGRLMIPNTGISLLIIPPDAIPRGKIYEIYLT	183
Qy	538	LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCE	597
Db	184	LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCE	243
Qy	598	GSEWDLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLLFAFPV	657
Db	244	GSEWDLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLLFAFPV	303
Qy	658	ACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPS	717
Db	304	ACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPS	363
Qy	718	SLWKSKLVSYSQEIPFYHIWNGTQRYLHCTFTLERVSPSTSIDLACKLWVWQVEGDGQSFS	777
Db	364	SLWKSKLVSYSQEIPFYHIWNGTQRYLHCTFTLERVSPSTSIDLACKLWVWQVEGDGQSFS	423
Qy	778	INFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPCCRGGADWRTLA	837
Db	424	INFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPCCRGGADWRTLA	483
Qy	838	QKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	897
Db	484	QKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	543
Qy	898	C 898	
Db	544	C 544	

RESULT 4

O08747

ID O08747 PRELIMINARY; PRT; 931 AA.
AC O08747;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Rostral cerebellar malformation protein.
GN UNC5H3 OR RCM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57B6/SJL;
RX MEDLINE=97271898; PubMed=9126743;
RA Ackerman S.L., Kozak L.P., Przyborski S.A., Rund L.A., Boyer B.B.,
RA Knowles B.B.;
RT "The mouse rostral cerebellar malformation gene encodes an UNC-5-like
protein.";
RL Nature 386:838-842(1997).
DR EMBL; U72634; AAB54103.1; -.
DR MGD; MGI:1095412; Unc5h3.
DR GO; GO:0005886; C:plasma membrane; IC.
DR GO; GO:0005042; F:netrin receptor activity; IDA.
DR GO; GO:0005515; F:protein binding; IDA.
DR GO; GO:0007420; P:brain development; IMP.
DR GO; GO:0030334; P:regulation of cell migration; IMP.
DR InterPro; IPR000488; Death.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00090; tsp_1; 2.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00408; IGc2; 1.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS50092; TSP1; 2.
KW Immunoglobulin domain.
SQ SEQUENCE 931 AA; 103062 MW; 8A5D951A4EECA179 CRC64;

Query Match 58.2%; Score 2787; DB 11; Length 931;
Best Local Similarity 57.3%; Pred. No. 9.7e-248;
Matches 522; Conservative 153; Mismatches 208; Indels 28; Gaps 9;

QY 9 PALLGIVLAAWRGSGAQQSA---TVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLVC 65
|| :|:| || || : | |: |||||:|||: ||||||| |||
Db 26 PAL--ALLSASGTGSAAQDDEFHHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYC 83

Qy	66	KAVPATQI FF KCN GEG VR QVD H VIER ST DGS SGL PT M E V R I N V S R Q Q V E K V F G L E E Y W C Q	125
Db	84	KAS PATQI YFK CN SEW VHQ KDH VV D E R V D E T S G L I V R E V S I E I S R Q Q V E E L F G P E D Y W C Q	143
Qy	126	CVAWSSSGTTKSQKAYIRIARL RKNF EQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLR	185
Db	144	CVAWSSAGTTKSRSRKAYVRIAYL RKT FEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK	203
Qy	186	NEDLVDP SLDP NVYITREHSLVVRQARLA DTANYTCVAKNIVARRRSASA A VIVYVNGGW	245
Db	204	NEDIIDPAEDRNFYITIDHNLI KQARLS DTANYTCVAKNIVAKRKSTTATVIVYVNGGW	263
Qy	246	STWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLC PVDGSWSPWS	305
Db	264	STWTEWSVCNSRCGRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIACTLC PVDGRWTSWS	323
Qy	306	KWSACGLDCTHWRSRECSDPAPRNGGEECQGTDL DTRNCTS DLCVHSASGPEDVALYVGL	365
Db	324	KWSTCGTECHWRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPDSDDVALYVG I	383
Qy	366	-IAVAVCLVLLLVLILVYCRKKEGLSDVADSI LTSGFQPVSIKPSKADNPHLLTIQP	424
Db	384	VIAVTVC LAITVVVALFVYRKNH RDFESDI IDSSALNGGFQPVN IKAARQD---LLAVPP	440
Qy	425	DLSTTTTYQGS LCP RQDG P S P K F Q L T NGHLLSPLGGGRHTLHHSS---PTSEAE F VS	480
Db	441	DLTSAAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLA EFSS	499
Qy	481	RLS---TQNYF-----RSLPRGT--SNMTYGT FNFLGGRLMIPNTGISL LIPPDAI	526
Db	500	KLSPQMTQSLLNEALNLKNQSLARQTDPSC TAFGT FNSLGGHLIIPNSGVSLI PAGAI	559
Qy	527	PRGKIYEIYLTLHKPEDVRLPIAGC QTLLS PIVSCGPGVLLTRPVILAMDHC GEPS PDS	586
Db	560	PQGRVYEMYVTVHRKENMRPPMEDSQ TLLTPV VSCGPGALLTRPVILTLHHCADP STED	619
Qy	587	WSLRLKKQSCEGSWEDVLHLGEEAPS HLYYCQ LEASAC YVFT EQLGRFALVGEALS VAAA	646
Db	620	WKIQ LKNQAVQGQWEDVVVGEENFTTPCYIQLDAEACHILTENLSTYALVGQSTTKAAA	679
Qy	647	KRLKLLL FAPVACTSLEYNIRVYCLHDTHDALKEV VQLEKQLGGQLIQEPRVLHFKDSYH	706
Db	680	KRLKLAIFG P LCCS SLEY SIRVYCLDDTQ DALKEV LQLERQMGGQ LLEEPKALHF KGSIH	739
Qy	707	NLRLSIHDV PSSLWKS KLLV SYQEIPFYHIWNGT QRYLHCTFTLERVSPSTS DLACKLWV	766
Db	740	NLRLSIHDIAHSIWLKS KLLAKYQEIPFYHIWSGSQRNLHCTFTLERLSLNTVELVCKLCV	799
Qy	767	WQVEGDGQSF SINFIN T KDT RFAELLA LESEAGVPA LVGPSAFKIPFLIRQKIISSLDPP	826
Db	800	RQVEGEGQI FQLNCTVSEEPTGIDLP LLDPASTITTVT GPSAFS I PLPIRQKLCS LDAP	859
Qy	827	CRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNIWEARHF PNGNLSQ LAAVAGLGQP	886
Db	860	QTRGH DWMLAHKLNLD RYLN YFATKSSPTGVILD LWEAQNF PDGNL SMLA AVLEEMGRH	919

QY 887 DAGLFTVSEAE 897
: : :| :
Db 920 ETVVSLAAEGQ 930

RESULT 5
Q8CD16
ID Q8CD16 PRELIMINARY; PRT; 950 AA.
AC Q8CD16;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Unc5 homolog.
GN UNC5H3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK031655; BAC27495.1; -.
DR PIR; PT0566; PT0566.
DR MGD; MGI:1095412; Unc5h3.
DR GO; GO:0005886; C:plasma membrane; IC.
DR GO; GO:0005042; F:netrin receptor activity; IDA.
DR GO; GO:0005515; F:protein binding; IDA.
DR GO; GO:0007420; P:brain development; IMP.
DR GO; GO:0030334; P:regulation of cell migration; IMP.
DR InterPro; IPR000488; Death.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00090; tsp_1; 2.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PRO1705; TSP1REPEAT.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGc2; 1.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS50092; TSP1; 2.
SQ SEQUENCE 950 AA; 105398 MW; 1E8FC74703351AF6 CRC64;

Query Match 57.8%; Score 2767.5; DB 11; Length 950;

Best Local Similarity 56.1%; Pred. No. 6.3e-246;
Matches 522; Conservative 153; Mismatches 208; Indels 47; Gaps 10;

Qy 9 PALLGIVLAAWLRGSGAQQSA---TVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLVC 65
| || : | : | || | : | | : | : | || | : | | | | | | | |
Db 26 PAL--ALLSASGTGSAAQDDEFHHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYC 83

Qy 66 KAVPATQI FFKCNGEWRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVGLEEYWCQ 125
| | | | : | | | | | | | | | | | | | | | | | | | | | | | |
Db 84 KASPAHQKDHVVDERVDESGLIVREVSEISRSQQVEELFGPEDYWCQ 143

Qy 126 CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWLR 185
| | | | : | | | : | | | | | | | | | | | | | | | | | | | | | |
Db 144 CVAWSSAGTTKSRKAYVRIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLK 203

Qy 186 NEDLVDPSDLPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVIVYVNGGW 245
| | | : | | | : | | | : | | | | | | | | | | | | | | | | | | | |
Db 204 NEDIIDPAEDRNFYITIDHNLIQARLSDTANYTCVAKNIVAKRKTATVIVYVNGGW 263

Qy 246 STWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDSWSPWS 305
| | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | |
Db 264 STWTEWSVCNSRCGRGYQKRTCTNPAPLNGGAFCEGQSVQKIACTTLCPVDGRWTSWS 323

Qy 306 KWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDTRNCTS DLCVH----- 351
| | | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 324 KWSTCGTECTHWRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQGFYPISTEHRPQN 383

Qy 352 -----SASGPEDVALYVGL-IAVAVCLVLLLVLILVYCRKKEGLDSDVADSSILTSGFQ 405
| | | | | | : | | | | | | | | | | | | | | | | | | | | | | | |
Db 384 EYGFSSAPSDDDVALYVGIVIAVTVCLAITVVVALFVYRKNHRDFESDIIDSSALNGGFQ 443

Qy 406 PVSIKPSKADNPPLLTIQPDLS TTTTYQGSLCPRQDGSPKFQLTNGHLLSPLGGRHT 465
| | : | | | | : | | | | | | | | | | | | | | | | | | | | | | |
Db 444 PVNIKAARQD--LLAVPPDLTSAAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIK 499

Qy 466 LHHSS----PTSEAEEFVSRSL---TQNYF-----RSLPRGT--SNMTYGTNFNLG 507
| | : | | | | : | | | | | | | | | | | | | | | | | | | | | | |
Db 500 VYNSSGAUTPQDDLAEFSSKLSPQMTQSLLNEALNLKNQSLARQTDPSCTAFGTFNSLG 559

Qy 508 GRLMIPNTGISLLIPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGV 567
| | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 560 GHЛИIPNSGVSLIIPAGAIPOGRVYEMYVTVHRKENMRPPMEDSQTLTPVSCGPPGAL 619

Qy 568 LTRPVILAMDHCGEPSDWSLRLKKQSCEGSWEDVLHLGEEAPS HLYYCQLEASACYVF 627
| | | | | | : | | | | | | | | | | | | | | | | | | | | | | | |
Db 620 LTRPVILTLHHCADPSTEDWKIQLKNQAVQGQWEDVVVGEENFTTPCYIQLDAEACHIL 679

Qy 628 TEQLGRFALVGEALSVAAKRLKLLL FAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQ 687
| | | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 680 TENLSTYALVGQSTTKAAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQ 739

Qy 688 LGGQLIQEPRVLHFKDSYHNLRLSIHDVPSLWKSLLVSYQEIPFYHIWNGTQRYLHCT 747
| | | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 740 MGGQLLEEPKALRFKGSIHNLRLSIHDIAHSLWKSLLAKYQEIPFYHIWSGSQRNLHCT 799

Qy 748 FTLERVSPSTS DLACKLWVWQVEGDGQFSINFNITKDTRFAELLALESEAGVPA LVGPS 807
| | | | : | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 800 FTLERLSLNTVELVCKLCVRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPASTITVTGPS 859
 Qy 808 AFKIPFLIRQKISSLDPCCRGA DWRTLAQKLHLDHLSFFASKPSPTAMILNLWEARH 867
 || || || || : || || | || || || || : || : || : || || : || : || : || : || :
 Db 860 AFSIPLPIRKLCSSLAPQT RGH DWRMLAHKLNLDRLNYFATKSSPTGVILDWEAQN 919
 Qy 868 FPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 897
 || : || || || : : | : : : | :
 Db 920 FPDGNL SMLAAVLEEMGRHETVVYLAEGQ 949

RESULT 6

Q7T2Z5

ID Q7T2Z5 PRELIMINARY; PRT; 931 AA.
 AC Q7T2Z5;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE UNC5-like protein 3.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Guan W., Condic M.L.;
 RT "Characterization of Netrin-1, Neogenin and cUNC-5H3 expression during chick dorsal root ganglia development.";
 RL Gene Expr. Patterns 3:369-373(2003).
 DR EMBL; AY187310; AAO67275.1; -.
 SQ SEQUENCE 931 AA; 102906 MW; 1E23A0D84F2E2C62 CRC64;

Query Match 57.6%; Score 2761; DB 13; Length 931;
 Best Local Similarity 57.0%; Pred. No. 2.4e-245;
 Matches 518; Conservative 151; Mismatches 212; Indels 28; Gaps 9;

Qy 9 PALLGIVLAAWL RGSGAQQS---ATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVL LVC 65
 ||| || | || || : | | : || || || : || || || || | |
 Db 26 PAL--AVLGASRPGSAAQDDDFHELPETFPSDPPPEPLPHFLIEPEEAYIVKNKPVNLYC 83
 Qy 66 KAVPATQIFFKCNGEW VRQVDHVIERSTDGSSGLPTMEVRINVS RQQVEKVGLEEYWCQ 125
 || || || || : || || | || || : | : || | || | : || || || : || | : || ||
 Db 84 KASPATQIYFKCNSEWHQKD HVVDERVDE TSGLIVCEVSIEISRQQVEELFGP EDYWCQ 143
 Qy 126 CVAWSSSGTTKSQKAYIRIARL RKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWL R 185
 || || || : || || : || || | || || || || || || || || : || | || || || || :
 Db 144 CVAWSSAGTTKS RKA YVRIAYLRKT FEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWL K 203
 Qy 186 NEDLVDP SLDPNVYITREHSLVVRQARLADT ANYTCVAKNIVARRSASA AVIVYVNGGW 245
 || :: || | | || : || : || : || || || || || || || || : || | || || || || :
 Db 204 NEEVIDPVEDRNFYITIDHNLI IKQARLSDT ANYTCVAKNIVAKRKSTTATVIVYVNGGW 263
 Qy 246 STWTEWSVCSASCGRGWQKRSRSTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWS 305
 || || || || | : || || : || : || || || || || || || || || || || || || | : ||
 Db 264 STWTEWSACNSRCGRGFQKRTRCTNPAPLNGGAFCEGQNVQKIACTTLC PVDGKWT SWS 323

Qy	306	KWSACGLDCTHWRSRECSDPAPRNGGEECQGTDDLDRNCTS DLCVHSASGPEDVALYVGL	365
Db	324	KWSTCGTECTHWRRRECTAPAKNGGKDCEGLVLQSKNCTDGLCMQAAPDSDDVALYVGI	383
Qy	366	-IAVAVCLVLLLLVILVYCRKKEGLSDVADSSILTSGFQPVS IKPSKADNP HLLTIQP	424
Db	384	VIAVIVCLAISVVVALFVYRKHNHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVPP	440
Qy	425	DLS TTTTYQGS LCP RQDGSPK FQLTNGHLLSPLGGGRHTLHHSS---PTSEAEFFVS	480
Db	441	DLTSAAMAMYRGPVYALHD-VSDKIPMTNSPILDPLPNLKIVK VYNTSGAVTPQD ELSDFSS	499
Qy	481	RLS---TQNYF-----RSLPRGT--SNMTYGT FNFLGGRLMIPNTGISLLIPPDAI	526
Db	500	KLSPQITQSLL ENETLNVKNQSLARQTDPSC TAFGTFN SLGGHLVIPNSGV SLLIPAGAV	559
Qy	527	PRGKIYEIYLTLHKPEDVRLPLAGCQTL LSPIVSCGPGVLLTRPVILAMDHC GEPS PDS	586
Db	560	PQGRVYEMYVTVHRKEGMRPPVEDSQ TLTPV VSCGPGALLTRPVVLT MHHCAEPNMDD	619
Qy	587	WSLRLKKQSCEGS WEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALS VAAA	646
Db	620	WQIQLKHQAGQGPWEDVVVVG EENFTT PCYIQLDPEACHILTE LSTYALVGQSITKAAA	679
Qy	647	KRLKLLL FAPVACTSLEYNIRVYCLHDTHDALKEV VQLEKQLGGQLIQEPRV LHF KDSYH	706
Db	680	KRLKLAIFGPLSCSSLEY SIRVYCLDDT QDALKEV LQLERQMGGQ LLEPKTLHFKGSTH	739
Qy	707	NLRLSIHDVPSLWKS KL LVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTS D LACK LWV	766
Db	740	NLRLSIHDIAHSLWKS KLP AKYQEIPFYHIWSGCQRNLHCTFTLERFS LNT LELVCKLCV	799
Qy	767	WQVEGDGQSFSINFN ITKDT RFAELL ALESEAGVPA LVGPSAFKIPFLIRQKIISSL DPP	826
Db	800	RQVEGEGQI FQLNC SVSEEPTGIDYPIMD SAGSITTIVGPNAFSIPLPIRQKLCSS LDAP	859
Qy	827	CRRGADWRTLAQKLHLD SHLSFFASKPSPTAMI LN IWEARHF PNGNLSQ LAAVAGLGQP	886
Db	860	QTRGH DWMLAHKL KLD RYL NYFATKSSPTGVILD LWEAQNF PDGNLSMLA AVLEEMGRH	919
Qy	887	DAGLFTVSE 895	
Db	920	: : :	
Qy	920	ETVVSLAAE 928	

RESULT 7

095185

ID 095185 PRELIMINARY; PRT; 931 AA.

AC 095185;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Transmembrane receptor UNC5C.

DEPARTMENT OF DEFENSE
TECHNICAL INFORMATION

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99000841; PubMed=9782087;
 RA Ackerman S.L., Knowles B.B.;
 RT "Cloning and mapping of the UNC5C gene to human chromosome 4q21-q23.";
 RL Genomics 52:205-208(1998).
 DR EMBL; AF055634; AAC67491.1; -.
 DR Genew; HGNC:12569; UNC5C.
 DR GO; GO:0005042; F:netrin receptor activity; TAS.
 DR GO; GO:0007411; P:axon guidance; TAS.
 DR GO; GO:0007420; P:brain development; TAS.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00090; tsp_1; 2.
 DR Pfam; PF00791; ZU5; 1.
 DR PRINTS; PRO1705; TSP1REPEAT.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00408; IGc2; 1.
 DR SMART; SM00209; TSP1; 2.
 DR SMART; SM00218; ZU5; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR PROSITE; PS50092; TSP1; 2.
 KW Immunoglobulin domain; Receptor.
 SQ SEQUENCE 931 AA; 103101 MW; EFD71122C98DABB8 CRC64;

Query Match 57.5%; Score 2755; DB 4; Length 931;
 Best Local Similarity 56.4%; Pred. No. 8.7e-245;
 Matches 514; Conservative 154; Mismatches 215; Indels 28; Gaps 9;

Qy	9	PALLGIVLAALRGSGAQQS---ATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLVLC	65
	: : : : :		
Db	26	PAL--ALLSASGTGSAAQDDDFHELPETFPSDPPEPLPHFLIEPSEEAYIVKNKPVNLYC	83
Qy	66	KAVPATQIIFFKCNGEWRQVDHVIERSTDGSSGLPTMEVRINVSQQVEKVFGLLEEWQC	125
	: :: : : :: :		
Db	84	KASPATQIYFKCNSEWHQKDHVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWQC	143
Qy	126	CVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEWL	185
	: : : :: :		
Db	144	CVAWSSAGTTKSRKAYVRIAYLRTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWL	203
Qy	186	NEDLVPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVIVYVNGGW	245
	: : : : : : : : : : : :		
Db	204	NEDIIDPVEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGW	263
Qy	246	STWTEWSVCSASCGRGWQKRSRSTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWS	305
	:: : : : : : : : : : : :		
Db	264	STWTEWSVCNSRCGRGYQKRTRCTNPAPLNGGAFCEGQSVQKIAC TTLCPVDGRWTPWS	323

RESULT 8

Q8JGT4

ID Q8JGT4 PRELIMINARY; PRT; 943 AA.

AC Q8JGT4;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBL rel. 22, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE UNC-5 receptor.

OS *Xenopus laevis* (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;

OC Xenopodinae; *Xenopus*.

OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Anderson R.B., Holt C.E.;
 RT "Expression of UNC-5 in the developing Xenopus visual system.";
 RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AY099459; AAM34486.1; -.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00090; tsp_1; 2.
 DR Pfam; PF00791; ZU5; 1.
 DR PRINTS; PRO1705; TSP1REPEAT.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00409; IG; 2.
 DR SMART; SM00408; IGc2; 1.
 DR SMART; SM00209; TSP1; 2.
 DR SMART; SM00218; ZU5; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR PROSITE; PS50092; TSP1; 2.
 KW Immunoglobulin domain; Receptor.
 SQ SEQUENCE 943 AA; 105083 MW; A024E24A7EDB6175 CRC64;

Query Match 55.2%; Score 2646.5; DB 13; Length 943;
 Best Local Similarity 53.0%; Pred. No. 9.3e-235;
 Matches 496; Conservative 163; Mismatches 229; Indels 47; Gaps 8;

Qy	10 ALLGIVLAAWLKG-----SGAQQSATVANPVPGANPDLLPHFLIVEPEDVYIVKNKPVL	62
	: : : : : : :	
Db	10 AALAAILVALILSCNFPSSTAGIEYSVDLPDSFPSAPAESLPHFLEPEDAYIVKNKPVE	69
Qy	63 LVCKAVPATQIIFFKCNGEWRQVDHIERSTDGSSGLPTMEVRINVSRRQQVEKVGLEY	122
	: : : : : :: :	
Db	70 LVCKANPATQIYFKCNGEWNQNNDHITKERVDDVTGLVVREVQIEVSRRQQVEELFGLEDY	129
Qy	123 WCQCVAWSSSGTTKSQAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVE	182
	: ::: : : : : : :	
Db	130 WCQCVAWSSAGTTKSRSYVRIAYLRKNFDQEPLGKEVALEQEALLQCRPPEGVPAAEVE	189
Qy	183 WLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVYVN	242
	: ::: : : ::: : : : : :	
Db	190 WLKNEEIIDPTKDTNFLITIDHNLIIKQARLSDTANYTCVSKNIVAKRRSTTATVIVFVN	249
Qy	243 GGWSTWTEWSVCSASCGRGWQKRSRSTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWS	302
	: : : : : :	
Db	250 GGWSSWTEWSPCNRCGHGWQKRRTCTNPAPLNGGTMCQGQYQKFACNTMCPVDGGWT	309
Qy	303 PWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTLDLDRNCTSILCVHS-----	352
	: : : : : :	

Db 310 EWSKWSACSTECTHWRRECNAPTPKNGGKDCSGMLLDKNCTDGLCMQNKRVLGETKSR 369
 Qy 353 -ASGPEDVALYVGL-IAAVAVCLVLLLVILVYCRKKEGLSDVADSS-ILTSGFQPVSI 409
 ||||| || :|: : ::|||: | |::| | | :|: ||| || || |||:
 Db 370 LLESTGDVALYAGLVVAIFIVIILLMAVGIVVYRRNCREFDTIDSSAALTGGFHPVNF 429
 Qy 410 KPSKADNPPLL--TIQPDLS||||TYQGSICPRQDGSPKFQLTNGHLLSPGGGRHTLH 467
 | |: || |: : |||||: | :||:: || : | :|| || || : ::
 Db 430 KTSRHDNSQLIHPAMQPDLTANAGIYRGNMYALQDS-ADKIPMTNSPLLDPLPNLKIKVY 488
 Qy 468 HSS-----PTSEAEVFVSRSLSTQN-----YFRSLPRGTSNMTYGTf 503
 :||| || : :: : | : :||| :||| |||
 Db 489 NSSTVGSSPGIHDGNNLLGKPTGTYPSDNNIMNARNKNMSMQHLLTLPRDSSNSVTGTF 548
 Qy 504 NFLGGRIMIPNTGISLLIIPPDAIIPRGKIYEIYLTLHKPEDVRLPLAGCQTLSPIVSCGP 563
 ||||| |||||:||||| |||||:|| ||:|| || :||| :|: | || | |||:|||:|||
 Db 549 GSLGGRLTFPNTGVSLIPOQAIPOQKYYEMYLMINKRENTVLPSEGTTQILSPIITCGP 608
 Qy 564 PGVLLTRPVILAMDHCGEPSDWSLRKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASA 623
 :||| :||| | : || : | | :||| || :||| :| ||| : ||| : ||| :|:
 Db 609 TGLLLCKPVILTVPHCADINTSDWILQLKTQSHQGNWEEVTLNEETLNTPCYCQLESHS 668
 Qy 624 CYVFTEQLGRFALVGEALSVAAKRLKLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ 683
 |: :||| | :| |||: | :| |||:| :|||: |||||||:|||: | | ||| :||:
 Db 669 CHTLLDQLGTYAFVGESYSRSAIKRLQLAI FAPMLCTSLEYNLKVYCVEDTPDALKEVLE 728
 Qy 684 LEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSLWKSKLLVSYQEIPFYHIWNGTQRY 743
 ||| ||| | :|||: | :||| | |||:||| | :|||: | |||:|||: | ||| :|||:
 Db 729 LEKTLGGYLVEEPKLLMFKDSYHNLRLSIHDIPHSLWRSKLMAKYQEIPFYHIWGSQRT 788
 Qy 744 LHCTFTLERVSPSTSDDLACKLWWVQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL 803
 |||||||| | :||| | :| |||: | |||:||| | :|||: | :|||: | :|||:
 Db 789 LHCTFTLERYSLAATELTCKICVRQVEGEGQIFQLHTLLEENVKSFDPECSQAENSVTTH 848
 Qy 804 VGPSAFKIPFLIRQKIISSLDPCCRGADWRTLAQKLHLDHLSFFASKPSPTAMILNLW 863
 :||| ||||| | :||| | ||| | ||| | :||| | :|||:| ||| :|||:
 Db 849 LGPYAFKIPFSIRQKICNSLDAPNSRGNDWRLLAQKLCMDRYLNYFATKASPTGVILDW 908
 Qy 864 EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 898
 || | :|||: | | :|: | :| | :| :| :|
 Db 909 EALHQDDGDLNTLASALEEMGKSEMMLVMATDGDC 943

RESULT 9

Q80Y85

ID Q80Y85 PRELIMINARY; PRT; 1008 AA.
 AC Q80Y85;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Unc5h2 protein (Fragment).
 GN UNC5H2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnurch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC048162; AAH48162.1; -.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000488; Death.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00090; tsp_1; 2.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00409; IG; 2.
DR SMART; SM00408; IGc2; 1.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS50092; TSP1; 2.
FT NON_TER 1 1
SQ SEQUENCE 1008 AA; 110438 MW; BCE5CA0EC537C130 CRC64;

Query Match 54.0%; Score 2585; DB 11; Length 1008;
Best Local Similarity 53.7%; Pred. No. 4.9e-229;
Matches 505; Conservative 151; Mismatches 235; Indels 50; Gaps 14;

Qy 1 MAVRPGLWPALLGIVLAAW-----LRG--SGAQQSATVANPVPGANPDLLPHFLVEPEDV 53
| | | : | | : | | | | | | : : | | : ||:||:||:||:
Db 75 MRARSGVRSALLALLCWDPPTSLAGVDSAGQ---VLPDSYPSAPAEQLPYFLLEPQDA 131

Qy 54 YIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHIERSTDGSSGLPTMEVRINVSRRQV 113
| | | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 132 YIVKNKPVELHCRAFPATQIYFKCNGEWSQNDHVTQESLDEATGLRVREVQIEVSRRQV 191

Qy 114 EKVGLEEYWCQCVAWSSSGTTKSQKAYIRIARLRNFEQEPLAKEVSLEQGIVLPCRPP 173
| | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 192 EELFGLEDYWCQCVAWSSSGTTKSRRAYIRIAYLRKNFDQEPLAKEVPLDHEVLLQCRPP 251

Qy 174 EGIPPAEVEWLRNEDLVDPSDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSA 233
| | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 252 EGVPVAEVEWLKNEDVIDPAQDTNFLLTIDHNLIIRQARLSDTANYTCVAKNIVAKRST 311

Qy 234 SAAVIVYVNGGWSTWTEWSVCASCGRGWQKRSRSTNPAPLNGGAFCEGQNVQKTACAT 293
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 312 TATVIVYVNGGWSSWAEWSPCSNRGCRGWQKRTRCTNPAPLNGGAFCEGQAFQKTACTT 371

Qy 294 LCPVDGSWSPWSKWSACGLDCTHWRSCRECSDPAPRNGGECCQGTDDTRNCTSDLCVHSA 353
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 372 VCPVDGAWTEWSKWSACSTECAHWRSCRECMAPPQNNGRDCSGTLLDSKNCTDGLCVLTL 431

Qy 354 SGPEDVALYVGL-TAVAVCLVLLLLVLILVYCRKKEGLSDVADSS-ILTSGFQPVSIKP 411
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 432 ETSGDVALYAGLVAVFVVAVLMAVGIVYRRNCRDFDTDITDSSAALTGGFHPVNFKT 491

Qy 412 SKADNPPLL--TIQPDLSSTTTTYQGSLCPRQDGSPKFQLTNGHLLSPLGGGRHTLHHS 469
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 492 ARPNNPQLLHPSAPPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLLDPLPSLKIKVYNS 550

Qy 470 S-----PTSEAEFSRLSTQNYFRS-----LPRGTSNMTY 500
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 551 STIGSGSGLADGADLLGVLPPTYPGDF-SRDTHFLHLRSASLGQHLLGLPRDPSSVS 609

Qy 501 GTFNFLGGRMLIPNTGISLLIPPDAPRGKIYEIYLTLHKPEDVRLPLA-GCQTLLSPIV 559
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 610 GTFGCLGGRLSLPGTVGSLLVNGAIPQGKFYDLYLHINKAEST-LPLSEGSQTVLSPSV 668

Qy 560 SCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSWEDVHLGEEAPSHLYYCQL 619
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 669 TCGPTGLLLCPVVLTVPHCAEVIAGDWIFQLKTQAHQGHWEVVTLDEETLNTPCYCQL 728

Qy 620 EASACYVFTEQLGRFALVGEALSVAAKRLKLLFAPVACTSLEYNIRVYCLHDTHDALK 679
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 729 EAKSCHILLDQLGYVFMGESYSRSAVKRLQLAIFAPALCTSLEYSLRVYCLEDPVALK 788

Qy 680 EVVQLEKQLGGQLIQEPRVLFHFKDSYHNLRLSIHDVPSLWKSKLLVSYQEIPFYHIWNG 739
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 789 EVLELERTLGGYLVEEPKPLLKD SYHNLRLSLHDIPH AHWRSKLLAKYQEIPFYHVWNG 848

Qy 740 TQRYLHCTFTLERVSPSTS DLACKLWVWQVEGDGQSF SINFNITKDTRFAELLALESEAG 799
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 849 SQRALHCTFTLERHSLASTEFTCKVCVRQVEGEGQIFQLHTTLA-ETPAGSLDALCSAPG 907

Qy 800 --VPALVGPSAFKIPFLIRQKIISSLPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTA 857

: :|| ||||| ||||| |||| | || ||| ||||| :| :||::||:| |||
Db 908 NAITTQLGPYAFKIPLSIRQKICSSLAPNSRGNDWRLLAQKLSMDRYLNYFATKASPTG 967
Qy 858 MILNLWEARHFPNGNLSQLAAVAGLGQPDAGLFTVSEAEC 898
:||:||| :||:||:||: :||: : : :|
Db 968 VILDLWEARQQDDGDILNSLASALEEMGKSEMLVAMATDGDC 1008

RESULT 10

Q8K1S3

ID Q8K1S3 PRELIMINARY; PRT; 945 AA.
AC Q8K1S3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Netrin receptor Unc5h2.
GN UNC5H2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Engelkamp D.;
RT "Cloning of three mouse unc-5 genes and their expression patterns at mid-gestation.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ487853; CAD32251.1; -.
DR MGD; MGI:894703; Unc5h2.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000488; Death.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00090; tsp_1; 2.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00409; IG; 2.
DR SMART; SM00408; IGc2; 1.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS50092; TSP1; 2.
KW Immunoglobulin domain; Receptor.
SQ SEQUENCE 945 AA; 103738 MW; 80E896F0F0E06012 CRC64;

Query Match 53.8%; Score 2578.5; DB 11; Length 945;
Best Local Similarity 53.2%; Pred. No. 1.8e-228;
Matches 506; Conservative 150; Mismatches 235; Indels 61; Gaps 15;

Qy 1 MAVRPGLWPALLGIVLAAW-----LRG--SGAQQSATVANPVPGANPDLLPHFLVEPEDV 53
| | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MRARSGVRSALLALLCWDPTPSLAGVDSAGQ--VLPDSYPSAPAEQLPYFLLEPQDA 57

Qy 54 YIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQQV 113
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 58 YIVKNKPVELHCRAFPATQIYFKCNGEWVSQNDHVTQESLDEATGLRVREVQIEVSRRQQV 117

Qy 114 EKVGLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKFQEPLAKEVSLEQGIVLPCRPP 173
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 118 EELFGLEDYWCQCVAWSSSGTTKSRRAYIRIAYLRKNFQDQEPLAKEVPLDHEVLLQCRPP 177

Qy 174 EGIPPAEVEWLNRNEDLVPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSA 233
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 178 EGVPVAEVEWLKNEDVIDPAQDTNFLLTIDHNLIIROARLSDTANYTCVAKNIVAKRST 237

Qy 234 SAAVIVYVNGGWSTWEWSVCASACGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACAT 293
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 238 TATVIVYVNGGWSSWAESPSCSNRCGRGWQKRTRCTNPAPLNGGAFCEGQAFQKTACTT 297

Qy 294 LCPVDGSWSPWSKWSACGLDCTHWRSCRECSDPAPRNGGECCQGTLDTRNCSDLCV--- 350
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 298 VCPVDGAWEWSKWSACSTECAHWRSCRECMAPPQNNGRDCSGTLLDSKNCTDGLCVLNQ 357

Qy 351 -----HSASGPEDVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLDSDVADSS-IL 400
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 358 RTLNDPKSHPLETSGDVALYAGLVVAVFVVAVLMAVGIVYRRNCRDFDTDITDSSAAL 417

Qy 401 TSGFQPVSIKPSKADNPHELL--TIQPDLSSTTTTYQGSLCPRQDGSPKFQLTNGHLLSP 458
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 418 TGGFHPVNFKTARPNNPQLLHPSAPPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLLDP 476

Qy 459 LGGGRHTLHHSS-----PTSEAEFFVSRLSTQNYFRS----- 490
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 477 LPSLKIKVYNSSSTIGSGGLADGADLLGVLPPTYPGDF-SRDTHFLHLRSASLGQSQHLL 535

Qy 491 -LPRGTSNMITYGTFNFLGGRLMIPNTGISLLIPPDATPRGKIYEIYLTLHKPEDVRLPLA 549
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 536 GLPRDPSSSVSGTFCGLGGRSLPPTGVSLVPGNAIPQGKFYDLYLHINKAEST-LPLS 594

Qy 550 -GCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSWEDVLHLGE 608
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 595 EGSQTVLSPSVCPTCGPTGLLCRPPVLTVPHCAEVIAQDWIFQLKTQAHQGHWEVVTLDE 654

Qy 609 EAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLFAPVACTSLEYNIRV 668
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 655 ETLNTPCYCQLEAKSCHILLDQLGTYVFMGESYSRSAVKRLQLAIFAPALCTSLEYSLRV 714

Qy 669 YCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSLWKSLLVSY 728
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 715 YCLEDTPVALKEVLELERTLGGYLVEEPKPLLKD SYHNLRLSLHDIPHAHWRSKLLAKY 774

Qy 729 QEIPFYHIWNGTQRYLHCTFTLERVSPSTS DLACKLWWQVEGDGQSFSINFNITKDTRF 788
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 775 QEIPFYHVWNGSQRALHCTFTLERHSLASTEFTCKVCVRQVEGEGQIFQLHTTLA-ETPA 833

Qy 789 AELLALESSEAG--VPALVGPSAFKIPFLIRQKISSLDPCCRRGADWRTLAQKLHLDHL 846
 | || | | : :|| |||| | ||| | ||| | ||| | ||| | ||| | :| :|
 Db 834 GSLDALCSAPGNAITTQLGPYAFKIPLSIRQKICSSLAPNSRGNDWRLLAQKLSMDRYL 893

 Qy 847 SFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 898
 :||:| ||| :||:|||| :||:|: ||:|: :|: : : :: :|
 Db 894 NYFATKASPTGVILDWEARQQDDGDLNSIASALEEMGKSEMLVAMATDGDC 945

RESULT 11

O08722

ID O08722 PRELIMINARY; PRT; 945 AA.
 AC O08722;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Transmembrane receptor UNC5H2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97271897; PubMed=9126742;
 RA Leonardo E.D., Hinck L., Masu M., Keino-Masu K., Ackerman S.L.,
 RA Tessier-Lavigne M.;
 RT "Vertebrate homologues of C. elegans UNC-5 are candidate netrin
 receptors.";
 RL Nature 386:833-838(1997).
 DR EMBL; U87306; AAB57679.1; -.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00090; tsp_1; 2.
 DR Pfam; PF00791; ZU5; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00408; IGc2; 1.
 DR SMART; SM00209; TSP1; 2.
 DR SMART; SM00218; ZU5; 1.
 DR PROSITE; PS50017; DEATH_DOMAIN; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR PROSITE; PS50092; TSP1; 2.
 KW Immunoglobulin domain; Receptor.
 SQ SEQUENCE 945 AA; 103520 MW; 6E9C2A262E560B9B CRC64;

Query Match 53.8%; Score 2578.5; DB 11; Length 945;
 Best Local Similarity 53.0%; Pred. No. 1.8e-228;
 Matches 509; Conservative 142; Mismatches 231; Indels 79; Gaps 17;

Qy	1	MAVRPGIWPALLGIVLAAW----LRG--SGAQQSATVANPVPGANPDLLPHFLVEPEDV	53
Db	1	MRARSGARGALLALLCWDPTPSLAGIDSGGQ--ALPDSFFSAPAEQLPHFLLEPEDA	57
Qy	54	YIVKNKPVLVCKAVPATQIFFKCNGEWRQVDHVIERSTDGSSGLPTMEVRINVSRRQV	113
Db	58	YIVKNKPVELHCRAFPATQIYFKCNGEWSQKGVTSQESLDEATGLRIREVQIEVSRRQV	117
Qy	114	EKVGLEEYWCQCVAWSSSGTTSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPP	173
Db	118	EELFGLEDYWCQCVAWSSSGTTSRRAYIRIAYLRKNFDQEPLAKEVPLDHEVLLQCRPP	177
Qy	174	EGIPPAEVEWLNRNEDLVDPSDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSA	233
Db	178	EGVPVAEVEWLKNEDVIDPAQDTNFLLTIDHNLIIHQARLSDTANYTCVAKNIVAKRST	237
Qy	234	SAAVIVYVNGGWSTWTEWSVCASCGRGWQKRSRSCTNPAPLNGGAFCEGQNQVKTACAT	293
Db	238	TATVIVYVNGGWSSWAESPCSNRCRGWQKRRTCTNPAPLNGGAFCEGQACQKTACTT	297
Qy	294	LCPVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEQCQGTDLDTRNCTS DLCV---	350
Db	298	VCPVDGAWTEWSKWSACSTECAHWRSCRMAPPQNGGRDCSGTLLSKNCTDGLCVLNQ	357
Qy	351	HSASGPE-----DVALYVGL-IAVAVCLVLLLVLILVYCRKKEGLDSDVADSS-IL	400
Db	358	RTLNDPKSRPLEPSGDVALYAGLVAVFVVLAVLMAVGIVYRRNCDFDTIDSSAAL	417
Qy	401	TSGFQPVSIKPSKADNPHELL--TIQPDLS TTTTYQGSLCPRQDGPSKFQLTNGHLLSP	458
Db	418	TGGFHPVNFKTARPSPNQPLLHPSAPPDLTASAGIYRGPVYALQDS-ADKI PMTNSPLLDP	476
Qy	459	L-----GGG-----RHTLHHSSPTSEAEFVS	480
Db	477	LPSLKIKVYDSSTIGSGAGLADGADLLGVLPPTYPGDFSRDTHFLHLRS-----A	527
Qy	481	RLSTQNYFRSLPRGTSNM TYGTFNFLGGRLMPNTGISLLIPPDAIPRGKIYEIYLTLHK	540
Db	528	SLGSQ-HLLGLPRDPSSSVSGTFGCLGGRLTIPGTGVSLVPGNAIPQGKFYDLYLRINK	586
Qy	541	PEDVRLPLA-GCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPS PDSWSLRLKKQSCEGS	599
Db	587	TEST-LPLSEGSQTVLSPSVTCGPTGLLCRPVVLTVPHCAEVIA GDWIFQLKTQAHQGH	645
Qy	600	WEDVLHLGEEAPSHLYYCQ LEASACYVFT EQLGRFA LVGEALS VAAKRLKLLFAPVAC	659
Db	646	WEEVVTLDEETLNTPCYCQLEAKSCHI LLDQLGTYVFTGESYSRS AVKRLQLAI FAPALC	705
Qy	660	TSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLFKDSYHNLRLSIHDV PSSL	719
Db	706	TSLEYSLRVYCLEDTPAALKEVLELERTLGGYLVEEPKTLLFKDSYHNLRLSLHDIPH AH	765
Qy	720	WKS KLLVSYQEIPFYHIWNGTQR YLHCTFTLERVSESTS DLACKLWVWQVEGDGQSFSIN	779
Db	766	WRSKLLAKYQEIPFYHWNGSQKALHCTFTLERHSI ASTEFTCKVCVRQVEGEGQIFQLH	825
Qy	780	FNITKDTRFAELLALESEAGVPA L--VGPSAFKIPFLIRQKIISSL DPPCRRGADWRTLA	837

```

          : :|   | || | | | :|| |||| | |||| :||| | || ||| |
Db      826 TTIA-ETPAGSLDALCSAPGNAATTQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRLLA 884

Qy      838 QKLHLDSHLSFFASKPSPTAMILNLWEARHFPGNQLAAVAGLGQPDAGLFTVSEAE 897
          ||| :| :||:||:| ||| :||:|||| :||:|: ||:|: :|: : : : :
Db      885 QKLSMDRYLNYFATKASPTGVILDWEARQQDDGDLNSLASALEEMGKSEMLVAMTTDGD 944

Qy      898 C 898
          |
Db      945 C 945

```

RESULT 12

Q9D398

ID Q9D398 PRELIMINARY; PRT; 945 AA.
AC Q9D398;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 6330415E02Rik protein.
GN UNC5H2 OR 6330415E02RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK018177; BAB31108.1; -.
DR MGD; MGI:894703; Unc5h2.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000488; Death.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.

||| |: ||| |||| :| ||:||:| |||:|| |::|| ::| | ||:
 Db 536 GLPRDPSSVSGTFCGLGRLSLPGTVSLLVNGAIPQGKFYDLYLHINKAEST-LPLS 594
 || |:||:| :||| |:|| |||:| :|| | | :|| |: |::||:| |
 Qy 550 -GCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSWEDVLHLGE 608
 || |:||:| :||| |:|| |||:| :|| | | :|| |: |::||:| |
 Db 595 EGSQLTVLSPTCGPTGLLLCPVVLTVPHCAEVIAQDWIFQLKTQAHQGHWEVVTLDE 654
 || : | |||| :||: :||| :||: | :| |||:| :|| | |||||:||:
 Qy 609 EAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLLFA PVACTSLEYNIRV 668
 || : | |||| :||: :||| :||: | :| |||:| :|| | |||||:||:
 Db 655 ETLNTPCYCQLEAKSCHILLDQLGSYVFMGESYSRSAVKRLQLAIFAPALCTSLEYSLRV 714
 ||| | | ||||:||:| ||| |:||: | | ||| | | |||:||:| :| :||| |
 Qy 669 YCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRSLSIHDVSSLWKSKLVSY 728
 ||| | | ||||:||:| ||| |:||: | | ||| | | |||:||:| :| :||| |
 Db 715 YCLEDTPVALKEVLELERTLGGYLVEEPKPLLFD SYHNLRSLSDI P H A W R S K L A K Y 774
 |||||:||:| ||| | | ||| | | :||: | | |||:|| | :| :|:
 Qy 729 QEIPFYHIWNGTQRYLHCTFTLERVSPSTS D LACKLWWVQ VEGDGQSFSINF NITKDTRF 788
 ||| | | ||||:||:| ||| | | ||| | | :||: | | |||:|| | :| :|:
 Db 775 QEIPFYHVWNGSQRALHCTFTLERHSI LASTEFTCKVCVRQVEGEGQI FQLHTTLA-ETPA 833
 ||| | | :| ||| | | ||| | | ||| | | |||:|| :| :|:
 Qy 789 AELLALESEAG--VPA LVGPSAFKIPFLIRQKIISSL DPPCRRGADWRTLAQKLHLD SHL 846
 ||| | | :| ||| | | ||| | | ||| | | ||| | | |||:|| :| :|:
 Db 834 GS LDALCSAPGNAITTQLGPYAFKIPLSIRQKICSSL DAPD S R G N D W R L L A Q K L S M D R Y L 893
 ||| | | :||:| | | :||:| ||| | | :||:| | :| :| :|:
 Qy 847 SFFASKPSPTAMI LNLWEARHF PNGNLS SQLAAVAGLGQPDAGLFTVSEAEC 898
 ||| | | :||:| | | :||:| ||| | | :||:| | :| :| :|:
 Db 894 NYFATKASPTGVILD IWEARQ QDDGDLNSL ASALEEMGKSEMLVAMATDGDC 945

RESULT 13

Q8IZJ1

ID Q8IZJ1 PRELIMINARY; PRT; 934 AA.
 AC Q8IZJ1;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Transmembrane receptor UNC5H2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22246081; PubMed=12359238;
 RA Komatsuzaki K., Dalvin S., Kinane T.B.;
 RT "Modulation of G(ialpha(2)) signaling by the axonal guidance molecule
 UNC5H2.";
 RL Biochem. Biophys. Res. Commun. 297:898-905 (2002).
 DR EMBL; AY126437; AAM95701.1; -.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR000906; ZU5.

Db 540 CLGGRLSIPGTGVSLVPNGAIPQGKFYEMYLLINKAEST-LPLSEGTQTVLSPSVTCGP 598
 Qy 564 PGVLLTRPVILAMDHCGEPSWLSRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASA 623
 |:|| ||||| | || | | :|| |: :| ||:| || : | || : | || |
 Db 599 TGLLLCPVILTMPCAEVSARDWIFQLKTQAHQGHWEVVTLDEETLNTPCYCQLEPRA 658
 Qy 624 CYVFTEQLGRFALVGEALSVAAKRLKLLFAPVACTSLEYNIRVYCLHDTHDALKEVQ 683
 |:: :||| : ||: | :| ||:| :||| | |||||:||| || | |||:
 Db 659 CHILLDQLGTYVFTGESYSRSAVKRLQLAVFAPALCTSLEYSLRVYCLEDTPVALKEVLE 718
 Qy 684 LEKQLGGQLIQEPRVLHFKDSYHNRLRLSIHDVPSLWKSLLVSYQEIPFYHIWNGTQRY 743
 ||: ||| |::||: | ||||||| |||:||:| :| :||| | ||||||| ||:|:
 Db 719 LERTLGGYLVEEPKPLMFKDSYHNRLRLSHDLPHAHWRSKLLAKYQEIPFYHIWSGSQKA 778
 Qy 744 LHCTFTLERVSPSTSDSLACKLWVWQVEGDGQFSINFNITKDTRFAELLALESEAG--VP 801
 | ||||||| | ::::| ||: | ||||:|| | :: :| | | | | |
 Db 779 LHCTFTLERHSLASTELTCKICVRQVEGEGQIFQLHTTLA-ETPAGSLDTLCAPGSTVT 837
 Qy 802 ALVGPSAFKIPFLIRQKIISSLDPCCRGADWRTLAQKLHLDHLSFFASKPSPTAMILN 861
 :|| | ||| | ||| :||| | || ||| ||| :| :||:||:| ||| :||:
 Db 838 TQLGPYAFKIPLSIRQKICNSLDAPNSRGNDWRMLAQKLSMDRYLNYFATKASPTGVILD 897
 Qy 862 LWEARHFPNGNLSQLAAAAGLGQPDAGLFTVSEAEC 898
 ||| | :||:| ||:| :||: : : :||:
 Db 898 LWEALQQDDGDNLNLASALEEMGKSEMLVAVATDGDC 934

RESULT 14

Q86SN3
 ID Q86SN3 PRELIMINARY; PRT; 945 AA.
 AC Q86SN3;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE P53-regulated receptor for death and life.
 GN P53RDL1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22533857; PubMed=12598906;
 RA Tanikawa C., Matsuda K., Fukuda S., Nakamura Y., Arakawa H.;
 RT "p53RDL1 regulates of p53-dependent apoptosis.";
 RL Nat. Cell Biol. 5:216-223(2003).
 DR EMBL; AB096256; BAC57998.1; -.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00531; death; 1.

DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00090; tsp_1; 2.
 DR Pfam; PF00791; ZU5; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00409; IG; 2.
 DR SMART; SM00408; IGc2; 1.
 DR SMART; SM00209; TSP1; 2.
 DR SMART; SM00218; ZU5; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR PROSITE; PS50092; TSP1; 2.
 KW Receptor.
 SQ SEQUENCE 945 AA; 103637 MW; 56064E335F323447 CRC64;

 Query Match 53.4%; Score 2558.5; DB 4; Length 945;
 Best Local Similarity 52.7%; Pred. No. 1.3e-226;
 Matches 501; Conservative 148; Mismatches 244; Indels 57; Gaps 15;

 Qy 1 MAVRPGIWPALLGIVLAAW---LRGSGAQQ-SATVANPVPGANPDLLPHFLVEPEDVYIV 56
 | | | | ||| : | | | | : | : | : | : ||:|| |:| | |||
 Db 1 MGARSGARGALLLALLLCWDPRLSQAGTDSGSEVLPDSFPSAPAEPFLPYFLQEPPQDAYIV 60

 Qy 57 KNKPVLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKV 116
 ||||| | |:| |||||:||||| | | | : | ::|| | |:| | ||| |::|:
 Db 61 KNKPVELRCRAFPATQIYFKCNGEWVSQNDHVTQEGLDEATGLRVREVQIEVSRRQQVEEL 120

 Qy 117 FGLEEWQCQCVAWSSSGTTKSQKAYIRIARLRLRNFEQEPLAKEVSLEQGIVLPCRPPEGI 176
 ||||:||||| |:| |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 Db 121 FGLEDYWCQCVAWSSAGTTKSRRAYVRIAYLRKNFDQEPLGKEVPLDHEVLLQCRPPEGV 180

 Qy 177 PPAEVEWLRNEDLVDPSDLPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAA 236
 | | | | |:| |:|:| | | :| | :|:|:|:|:|:|:|:|:|:|:|:|:
 Db 181 PVAEVEWLKNEDVIDPTQDTNFLLTIDHNLIIQARLSDTANYTCVAKNIVAKRRSTTAT 240

 Qy 237 VIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCP 296
 ||| | | | |:| | | | | | | | | | | | | | | | | | | | | | | |:
 Db 241 VIVYVNGGWSSWAESWPCSNRCGRGWQKRRTCTNPAPLNGGAFCEGQAFQKTACTTICP 300

 Qy 297 VDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTLDTRNCTS DLCVHS---- 352
 ||:|: | | | | | :| | | | | | | | | | | | | | | | | | | | |:
 Db 301 VDGAWTEWSKWSACSTECAHWRSRECMAPPQNNGRDCSGTL LDSKNCTDGLCMQNKCTL 360

 Qy 353 -----ASGPEDVALYVGL-IAVAVCLVLLLVLILVYCRKKEGLDSDVADSS-ILT 401
 ||| | | | | | :| | | :| | | :| | | | | | | | | | | | | | |:
 Db 361 SDPNSHLLEASG--DAALYAGLVVAIFVVVAILMAVGVVYRRNCRDFDTDITDSSAALT 418

 Qy 402 SGFQPVSIKPSKADNPHELL--TIQPDLSTTTTYQGSCLPRQDGPSPKFQLTNGHLLSPL 459
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |:
 Db 419 GGFHPVNFKTARPSNPQLLHPSPVPPDLTASAGIYRGPVYALQDS-TDKIPMTNSPLLDPL 477

 Qy 460 GGGRHTLHHSSPT-----SEAEFFVSRLSTQNY-----FRS-----L 491
 : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |:
 Db 478 PSLKVKVYSSSTTGSGPGLADGADLLGVLPPIGYPSDFARDTHFLHLRSASLGSQQLLGL 537

 Qy 492 PRGTSNMITYGTFNFLGGRLMIPNTGISLIPPAI PRGKIYEIYLTLHKPEDVRLPLA-G 550
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |:
 Db 538 PRDPGSSVSGTFGCLGGRLSIPGTGVSLVPNGAIPOGKFYEMYLLINKAEST-LPLSEG 596

RESULT 15

Q8K1S2

ID Q8K1S2 PRELIMINARY; PRT; 956 AA.

AC Q8K1S2;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Netrin receptor Unc5h4.

GN UNC5H4.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Engelkamp D.;

RT "Cloning of three mouse unc-5 genes and their expression patterns at mid-gestation.";

RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AJ487854; CAD32252.1; -.

DR MGD; MGI:2389364; Unc5h4.

DR GO; GO:0004872; F:receptor activity; IEA.

DR GO; GO:0004867; F:serine protease inhibitor activity; IEA.

DR GO; GO:0007165; P:signal transduction; IEA.

DR InterPro; IPR000488; Death.

DR InterPro; IPR003599; Ig.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003598; Ig_c2.

DR InterPro; IPR000215; Serpin.

DR InterPro; IPR000884; TSP1.

DR InterPro; IPR008085; TSP_1.

DR InterPro; IPR000906; ZU5.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF00090; tsp_1; 2.
 DR Pfam; PF00791; ZU5; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00408; IGc2; 1.
 DR SMART; SM00209; TSP1; 2.
 DR SMART; SM00218; ZU5; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 DR PROSITE; PS50092; TSP1; 2.
 KW Immunoglobulin domain; Receptor.
 SQ SEQUENCE 956 AA; 106351 MW; DFDF07839C10C68D CRC64;

Query Match 45.9%; Score 2200; DB 11; Length 956;
 Best Local Similarity 45.5%; Pred. No. 1.6e-193;
 Matches 431; Conservative 159; Mismatches 280; Indels 78; Gaps 15;

Qy	8 WPALLGIVLAAWLRGS---GAQQSATVANPVGANPDLLPHFLVEPEDVYIVKNKPVLL	63
	: : : : : : : : :	
Db	15 WLPWLGLFF--WAAGAAAARGADGSEILPDPSA-PGTLPHFIEEPEDAYIIKSNPIAL	71
Qy	64 VCKAVPATQIFFKCNGEWRQVDHIERSTDGSSGLPTMEVRINVSQQVEKVFGLLEYW	123
	: : :	
Db	72 RCKARPAMQIFFKCNGEWRHQNEHVSEESLDDESSGLKVREVFINVTRQQVEDFHGPEDYW	131
Qy	124 CQCVAWSSTTSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAEVEW	183
	: : : : : : :	
Db	132 CQCVAWSHLGTSKSRKASVRIAYLRKNFEQDPQGREVPIEGMIVLHCRPPEGVPAAEVEW	191
Qy	184 LRNEDLVDPSDLPNVYITREHSILVVRQARLADTANYTCVAKNIVARRRSASAIVYVNG	243
	: : : : : : : : : : : : :	
Db	192 LKNEEPIDSEQDENIDTRADHNLIIRQARLSDSGNYTCMAANIVAKRRSLSATVVVYVNG	251
Qy	244 GWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSP	303
	: : : : : :	
Db	252 GWSSWTEWSACNVRCGRGWQKRSRTCTNPAPLNGGAFCEGMSVQKITCTALCPVDGSWEV	311
Qy	304 WSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDRNCTS DLCV-----	350
	: : : : : : : : :	
Db	312 WSEWSVCSPCEHLIRECTAPPNGKFCEGLSQESENCTDGLCILDKKPLHEIKPQR	371
Qy	351 ---HSASGPEDVALYVGLIAVAVCLVLLLVLILVYCRKKEGLDSDVADSSILTSGFQPV	407
	: : : :	
Db	372 WSRRGIENASDIALYSGL-GAAVVAVAVLVIIGVTLYRRSHSDYGVVIDSSALTGGFQTF	430
Qy	408 SIKPSKADNPPLL--TIQPDLS TYQGSLCPRQDGSPKFQLTNGHLLSPLG----	460
	: : : :	
Db	431 NFKTVRQGNSSLNPAMQPDL-TVSRTYSGPIC-LQD-PLDKELMTESSLFNPLSDIKVK	487
Qy	461 -----GGRH-----TLHHSSPTSEAEVFVSRNSTQNYFR	489
	: :	
Db	488 VQSSFMVSLGVSERAEYHGKNHSGTFPHGNRGFSTIHPRNKT---PYIQNLS-----	537

Qy 490 SLPRGTSNMTYGTNFLGGRLMPINTGISLLIPPAIPRGKIEIYTLHLKPEDVRLPLA 549
| | | | | | | | : | | | | | | | : | | | | | |
Db 538 SLPTRTELRTTGVFGHGGRLVMPNTGVSLIPHGAIPPEENSWEIYMSINQGEP-SLQSD 596

Qy 550 GCQTLLSPLIVSCGPPGVLLTRPVLAMDHCGEPSDWSLRLKKQSCEGSWEDVLHLGEE 609
| : | | | | : | | | : | | | : | | | : | | | : | | | : | | | : | | |
Db 597 GSEVLLSPEVTGPPDMLVTTPFALTIPHCADVSSEHWNILHKRTQQGKWEEVMSVEDE 656

Qy 610 APSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAKRLKLLLFAFPVACTSLEYNIRVY 669
: | | | | : | | | : | | | | : | | | : | | | : | | | : | | |
Db 657 STS--CYCLLDPFACHVLLDSFGTYALTGEPI TDCAVKQLKVAVFGCMSCNSLDYNLRVY 714

Qy 670 CLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSDYHNLRLSIHDVSSLWKS KLLVSYQ 729
| : | | | : | | | | : | | | : | | | : | | | : | | |
Db 715 CVDNTPCAFQEVISDERHQGGQLLEEPKLLHFKGNTFSLQSVLDIPPFLWRIKPFTACQ 774

Qy 730 EIPFYHIWNGTQRYLHCTFTLERVSPSTS DLA CKLWVWQVEGDGQSFSINFNITKDTRFA 789
| : | | : | : | | | | : | | : | | : | | : | | : | | : |
Db 775 EVPFSRVWSSNRQPLHCAFSLERYTPTTTQLSCKICIRQLKGHEQILQVQTSILESERET 834

Qy 790 ELLALESEAGVPALVGPSAFKIPFLIRQKIISLDPPCRRGADWRTL A QKLHLD SHLSFF 849
: | | | | | | | | : | | | | | | | | | | | | | | | | |
Db 835 ITFFAQEDSTFPAQTGPKA FKI PYSIRQRICATFDTPNAKGKD WQMLAQKNSINRNLSYF 894

Qy 850 ASKPSPTAMILNLWEARHFPNGNLSQLAAA VAGLGQPDAGLFTVSEAE 897
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 895 ATQSSPSAVILNLWEARHQDGDLDLACALEEIGRTHTKLSNITEPQ 942

Search completed: July 6, 2004, 14:34:59
Job time : 53 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2004, 14:20:26 ; Search time 17 Seconds
(without alignments)
2750.529 Million cell updates/sec

Title: US-10-624-932-2

Perfect score: 4791

Sequence: 1 MAVRPGLWPALLGIVLAAWL.....AVAGLGQPDAGLFTVSEAE 898

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			Description
		Match	Length	DB	
1	298.5	6.2	1584	1 BAI1_HUMAN	O14514 homo sapien
2	296.5	6.2	1172	1 TSP2_HUMAN	P35442 homo sapien
3	293	6.1	1074	1 SM5A_HUMAN	Q13591 homo sapien
4	293	6.1	1172	1 TSP2_MOUSE	Q03350 mus musculu
5	291.5	6.1	1170	1 TSP2_BOVIN	Q95116 bos taurus
6	291	6.1	1077	1 SM5A_MOUSE	Q62217 mus musculu
7	290	6.1	1173	1 TSP1_XENLA	P35448 xenopus lae
8	276	5.8	1093	1 SM5B_HUMAN	Q9p283 homo sapien
9	275.5	5.8	1093	1 SM5B_MOUSE	Q60519 mus musculu
10	275	5.7	1522	1 BAI3_HUMAN	O60242 homo sapien
11	274.5	5.7	1572	1 BAI2_HUMAN	O60241 homo sapien
12	270.5	5.6	1170	1 TSP1_MOUSE	P35441 mus musculu
13	268.5	5.6	1170	1 TSP1_HUMAN	P07996 homo sapien
14	265.5	5.5	1170	1 TSP1_BOVIN	Q28178 bos taurus
15	263	5.5	1178	1 TSP2_CHICK	P35440 gallus gall
16	243	5.1	469	1 PROP_HUMAN	P27918 homo sapien
17	240	5.0	470	1 PROP_CAVPO	Q64181 cavia porce

18	229	4.8	437	1	PROP_MOUSE	P11680	mus musculu
19	201	4.2	867	1	SSPO_BOVIN	P98167	bos taurus
20	199	4.2	1266	1	NGCA_CHICK	Q03696	gallus gall
21	181	3.8	1736	1	ZO1_HUMAN	Q07157	homo sapien
22	178	3.7	1745	1	ZO1_MOUSE	P39447	mus musculu
23	177	3.7	587	1	CO8B_ONCMY	Q90x85	oncorhynchus
24	175.5	3.7	905	1	ATS8_MOUSE	P57110	mus musculu
25	173.5	3.6	630	1	ATS4_RAT	Q9esp7	rattus norv
26	172.5	3.6	1223	1	AT14_HUMAN	Q8wxs8	homo sapien
27	172	3.6	837	1	ATS4_HUMAN	O75173	homo sapien
28	170.5	3.6	1077	1	AT10_HUMAN	Q9h324	homo sapien
29	170.5	3.6	1224	1	AT16_HUMAN	Q8te57	homo sapien
30	169.5	3.5	890	1	ATS8_HUMAN	Q9up79	homo sapien
31	169.5	3.5	934	1	CO6_HUMAN	P13671	homo sapien
32	168	3.5	860	1	ATS6_HUMAN	Q9ukp5	homo sapien
33	168	3.5	1095	1	AT17_HUMAN	Q8te56	homo sapien
34	165.5	3.5	1205	1	ATS3_HUMAN	O15072	homo sapien
35	161.5	3.4	1906	1	AT20_MOUSE	P59511	mus musculu
36	160	3.3	930	1	ATS5_HUMAN	Q9una0	homo sapien
37	160	3.3	930	1	ATS5_MOUSE	Q9r001	mus musculu
38	159.5	3.3	967	1	ATS1_RAT	Q9wuq1	rattus norv
39	159.5	3.3	968	1	ATS1_MOUSE	P97857	mus musculu
40	158.5	3.3	997	1	ATS7_HUMAN	Q9ukp4	homo sapien
41	157	3.3	967	1	ATS1_HUMAN	Q9uh18	homo sapien
42	156.5	3.3	562	1	AT15_MOUSE	P59384	mus musculu
43	156.5	3.3	807	1	FSPO_RAT	P35446	rattus norv
44	155.5	3.2	1911	1	AT20_HUMAN	P59510	homo sapien
45	155	3.2	584	1	CO8A_HUMAN	P07357	homo sapien

ALIGNMENTS

RESULT 1

BAI1_HUMAN

ID BAI1_HUMAN STANDARD; PRT; 1584 AA.

AC O14514;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Brain-specific angiogenesis inhibitor 1 precursor.

GN BAI1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Fetal brain;

RX MEDLINE=98054121; PubMed=9393972;

RA Nishimori H., Shiratsuchi T., Urano T., Kimura Y., Kiyono K.,

RA Tatsumi K., Yoshida S., Ono M., Kuwano M., Nakamura Y., Tokino T.;

RT "A novel brain-specific p53-target gene, BAI1, containing

RT thrombospondin type 1 repeats inhibits experimental angiogenesis.";

RL Oncogene 15:2145-2150(1997).

RN [2]

RP INTERACTION WITH BAP1.

RX MEDLINE=98321173; PubMed=9647739;
RA Shiratsuchi T., Futamura M., Oda K., Nishimori H., Nakamura Y.,
RA Tokino T.;
RT "Cloning and characterization of BAI-associated protein 1: a PDZ
RT domain-containing protein that interacts with BAI1.";
RL Biochem. Biophys. Res. Commun. 247:597-604(1998).
CC -!- FUNCTION: LIKELY TO BE A POTENT INHIBITOR OF ANGIOGENESIS IN
CC BRAIN AND MAY PLAY A SIGNIFICANT ROLE AS A MEDIATOR OF THE P53
CC SIGNAL IN SUPPRESSION OF GLIOBLASTOMA. MAY FUNCTION IN CELL
CC ADHESION AND SIGNAL TRANSDUCTION IN THE BRAIN.
CC -!- SUBUNIT: INTERACTS WITH BAP1.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. LIKELY TO BE
CC CONCENTRATED AT CELL-CELL ADHESION SITES.
CC -!- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN BRAIN. REDUCED OR NO
CC EXPRESSION IS OBSERVED IN SOME GLIOBLASTOMA CELL LINES AND CANCER
CC TISSUES.
CC -!- INDUCTION: By p53.
CC -!- DOMAIN: THE TSP1 REPEATS INHIBIT IN VIVO ANGIOGENESIS IN RAT
CC CORNEA INDUCED BY BFGF.
CC -!- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.
CC -!- SIMILARITY: Contains 5 TSP type-1 domains.
CC -!- SIMILARITY: Contains 1 GPS domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB005297; BAA23647.1; -.
DR PIR; T00026; T00026.
DR Genew; HGNC:943; BAI1.
DR MIM; 602682; -.
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0005911; C:intercellular junction; TAS.
DR GO; GO:0005515; F:protein binding; TAS.
DR GO; GO:0007409; P:axonogenesis; TAS.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
DR GO; GO:0007422; P:peripheral nervous system development; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR000832; GPCR_secretin.
DR InterPro; IPR001879; hormn_receptor.
DR InterPro; IPR000203; PKD_cys_rich.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00002; 7tm_2; 1.
DR Pfam; PF01825; GPS; 1.
DR Pfam; PF02793; HRM; 1.
DR Pfam; PF00090; tsp_1; 5.
DR SMART; SM00303; GPS; 1.
DR SMART; SM00008; HormR; 1.
DR SMART; SM00209; TSP1; 5.
DR PROSITE; PS50221; GPS; 1.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; FALSE_NEG.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; FALSE_NEG.

DR PROSITE; PS50227; G_PROTEIN_RECEP_F2_3; 1.
 DR PROSITE; PS50261; G_PROTEIN_RECEP_F2_4; 1.
 DR PROSITE; PS50092; TSP1; 5.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Repeat; Cell adhesion.
 FT SIGNAL 1 30 POTENTIAL.
 FT CHAIN 31 1584 BRAIN-SPECIFIC ANGIOGENESIS INHIBITOR 1.
 FT DOMAIN 31 948 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 949 969 1 (POTENTIAL).
 FT DOMAIN 970 980 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 981 1001 2 (POTENTIAL).
 FT DOMAIN 1002 1008 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1009 1029 3 (POTENTIAL).
 FT DOMAIN 1030 1052 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 1053 1073 4 (POTENTIAL).
 FT DOMAIN 1074 1093 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1094 1114 5 (POTENTIAL).
 FT DOMAIN 1115 1136 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 1137 1157 6 (POTENTIAL).
 FT DOMAIN 1158 1166 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1167 1187 7 (POTENTIAL).
 FT DOMAIN 1188 1584 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 261 315 TSP TYPE-1 1.
 FT DOMAIN 354 407 TSP TYPE-1 2.
 FT DOMAIN 409 462 TSP TYPE-1 3.
 FT DOMAIN 467 520 TSP TYPE-1 4.
 FT DOMAIN 522 575 TSP TYPE-1 5.
 FT DOMAIN 881 938 GPS.
 FT DOMAIN 1411 1422 POLY-PRO.
 FT DOMAIN 1425 1430 POLY-PRO.
 FT SITE 231 233 CELL ATTACHMENT SITE (POTENTIAL).
 FT DOMAIN 1365 1584 NECESSARY FOR INTERACTION WITH BAP1.
 FT DOMAIN 1581 1584 INDISPENSABLE FOR INTERACTION WITH BAP1.
 FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 692 692 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 844 844 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 877 877 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 881 881 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 1584 AA: 173531 MW: DEA8F28C77874513 CRC64:

Query Match 6.2%; Score 298.5; DB 1; Length 1584;
Best Local Similarity 33.5%; Pred. No. 2.7e-14;
Matched 52, Total 125, Missed 73, False 21, Total 83, Score 16

Ov 184 LRNEDLVDEPSLDPNVYITREHSLVVROARLADTANYTCVAKNIVARRSASAALIVYVNG 243

Ph 361 SPWCHSSSTGCCNWTB TPDCHSSSYSTGCGBLPQBLGNNSAUGPHC 410

Db 411 AWDEWSPWSLCSSTCGRGFRDRTRTCR--PPQFGGNPCEGPEKQTFCNIALCPGRAVDG 468

QY 300 SWSPWSKWSACGLDCT---HWRSCSDPAPRNGGEECQGTLDTRNCTS DLC 349
:|: || |||| |: |:|||: |: || |||| :||:| |
Db 469 NWNEWSSWSACSASC SQGRQQRTRECNGPS--YGGAE CQGHVETRDCFLQOC 519

RESULT 2

TSP2_HUMAN

ID TSP2_HUMAN STANDARD; PRT; 1172 AA.
AC P35442;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thrombospondin 2 precursor.
GN THBS2 OR TSP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94010892; PubMed=8406456;
RA Labell T.L., Byers P.H.;
RT "Sequence and characterization of the complete human thrombospondin 2
cDNA: potential regulatory role for the 3' untranslated region.";
RL Genomics 17:225-229(1993).
RN [2]
RP SEQUENCE OF 560-1172 FROM N.A.
RC TISSUE=Fibroblast;
RX MEDLINE=92217961; PubMed=1559694;
RA Labell T.L., McGookey Milewicz D.J., Disteche C.M., Byers P.H.;
RT "Thrombospondin II: partial cDNA sequence, chromosome location, and
expression of a second member of the thrombospondin gene family in
humans.";
RL Genomics 12:421-429(1992).
RN [3]
RP THROMBOSPONDIN REPEATS DISULFIDE BONDS.
RX MEDLINE=21588233; PubMed=11590138;
RA Misener T.M., Hahr A.J., Harms A.C., Annis D.S., Mosher D.F.;
RT "Disulfide connectivity of recombinant C-terminal region of human
thrombospondin 2.";
RL J. Biol. Chem. 276:45882-45887(2001).
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
laminin and type V collagen.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; L12350; AAA03703.1; -.
 DR EMBL; M81339; -; NOT_ANNOTATED_CDS.
 DR PIR; A47379; TSHUP2.
 DR HSSP; P00740; 1EDM.
 DR Genew; HGNC:11786; THBS2.
 DR MIM; 188061; -.
 DR GO; GO:0008201; F:heparin binding; TAS.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR003367; tsp_3.
 DR InterPro; IPR008859; TSPC.
 DR InterPro; IPR003129; TSPN.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00090; tsp_1; 3.
 DR Pfam; PF02412; tsp_3; 13.
 DR Pfam; PF05735; TSPC; 1.
 DR Pfam; PF02210; TSPN; 1.
 DR Pfam; PF00093; vwc; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00181; EGF; 3.
 DR SMART; SM00209; TSP1; 3.
 DR SMART; SM00210; TSPN; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS50026; EGF_3; 2.
 DR PROSITE; PS50092; TSPI; 3.
 DR PROSITE; PS01208; VWFC_1; 1.
 DR PROSITE; PS50184; VWFC_2; 1.
 KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
 KW EGF-like domain; Signal.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 1172 THROMBOSPONDIN 2.
 FT DOMAIN 19 215 TSP N-TERMINAL.
 FT DOMAIN 19 232 HEPARIN-BINDING (POTENTIAL).
 FT DOMAIN 318 375 VWFC.
 FT DOMAIN 381 431 TSP TYPE-1 1.
 FT DOMAIN 437 492 TSP TYPE-1 2.
 FT DOMAIN 494 549 TSP TYPE-1 3.
 FT DOMAIN 549 589 EGF-LIKE 1.
 FT DOMAIN 590 647 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 648 692 EGF-LIKE 3.
 FT DOMAIN 725 760 TSP TYPE-3 1.
 FT DOMAIN 761 783 TSP TYPE-3 2.
 FT DOMAIN 784 819 TSP TYPE-3 3.
 FT DOMAIN 820 842 TSP TYPE-3 4.
 FT DOMAIN 843 880 TSP TYPE-3 5.
 FT DOMAIN 881 916 TSP TYPE-3 6.
 FT DOMAIN 917 952 TSP TYPE-3 7.

Db 565 -----SSFPDGS-WSCGFCPVGFLGNTHCEDLDECALVPDICFSTSKVPRCVNTQP 615
Qy 435 GSLC----PRQDGPS 446
| | || | |
Db 616 GFHCLPCPPRYRGNQP 631

RESULT 3

SM5A_HUMAN

ID SM5A_HUMAN STANDARD; PRT; 1074 AA.
AC Q13591; O60408;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Semaphorin 5A precursor (Semaphorin F) (Sema F).
GN SEMA5A OR SEMAF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98125554; PubMed=9464278;
RA Simmons A.D., Puschel A.W., McPherson J.D., Overhauser J., Lovett M.;
RT "Molecular cloning and mapping of human semaphorin F from the Cri-du-
RT chat candidate interval.";
RL Biochem. Biophys. Res. Commun. 242:685-691(1998).
RN [2]
RP SEQUENCE OF 1-494 FROM N.A.
RA Kalicki J., Harmon G.;
RL Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
CC --!- FUNCTION: May act as positive axonal guidance cues.
CC --!- SUBCELLULAR LOCATION: Type I membrane protein.
CC --!- SIMILARITY: Belongs to the semaphorin family.
CC --!- SIMILARITY: Contains 1 Sema domain.
CC --!- SIMILARITY: Contains 7 TSP type-1 domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U52840; AAC09473.1; -.
DR EMBL; AC004615; AAC14668.1; -.
DR PIR; JC5928; JC5928.
DR Genew; HGNC:10736; SEMA5A.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR GO; GO:0007267; P:cell-cell signaling; TAS.
DR GO; GO:0007399; P:neurogenesis; TAS.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002165; Plexin_repeat.
DR InterPro; IPR001627; Sema.
DR InterPro; IPR000884; TSP1.

DR InterPro; IPR008085; TSP_1.
 DR Pfam; PF01437; PSI; 1.
 DR Pfam; PF01403; Sema; 1.
 DR Pfam; PF00090; tsp_1; 6.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR SMART; SM00209; TSP1; 6.
 DR PROSITE; PS50092; TSP1; 6.
 KW Signal; Transmembrane; Repeat; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 1074 SEMAPHORIN 5A.
 FT DOMAIN 23 968 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 969 989 POTENTIAL.
 FT DOMAIN 990 1074 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 226 507 SEMA.
 FT DOMAIN 540 593 TSP TYPE-1 1.
 FT DOMAIN 595 651 TSP TYPE-1 2.
 FT DOMAIN 653 702 TSP TYPE-1 3.
 FT DOMAIN 707 765 TSP TYPE-1 4.
 FT DOMAIN 784 839 TSP TYPE-1 5.
 FT DOMAIN 841 896 TSP TYPE-1 6.
 FT DOMAIN 897 944 TSP TYPE-1 7.
 FT CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 323 323 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 367 367 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 717 717 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 933 933 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 56 56 A -> V (IN REF. 2).
 FT CONFLICT 149 149 A -> T (IN REF. 2).
 FT CONFLICT 382 382 V -> M (IN REF. 2).
 FT CONFLICT 494 494 S -> R (IN REF. 2).
 SO SEQUENCE 1074 AA: 120570 MW: EE3DB763CBF29407 CRC64:

```

Query Match           6.1%; Score 293; DB 1; Length 1074;
Best Local Similarity 45.8%; Pred. No. 4.1e-14;
Matches 54; Conservative 11; Mismatches 49; Indels 4; Gaps 2;

Y 241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTACATL-CPVDG 299
| | | | | | | | | | | | : | | | | | | | | | | | | ::: | | | | | |
D 783 VNGAWSAWTWSQCSRDCSRGIRNRKRCVNNPEPKYGGMPCLGPSLEYQECNTLPCPVDG 842

Y 300 SWSPWSKWSACGLDC---THWRSRECSDPAPRNGGEECQGTDLDTRNCTS DLCVHSAS 354
| | | | : | | | : | : | | | | | | : | | | : | | | | | |
D 843 VWSCWSWPWTKCSATCGGGHYMRTRSCSNPAPAYGGDICLGLHTEALCNTQPCPESWS 900

```

RESULT 4
TSP2_MOUSE
ID TSP2_MOUSE STANDARD; PRT; 1172 AA.

AC Q03350;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thrombospondin 2 precursor.
GN THBS2 OR TSP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92147683; PubMed=1371115;
RA Laherty C.D., O'Rourke K., Wolf F.W., Katz R., Seldin M.F.,
RA Dixit V.M.;
RT "Characterization of mouse thrombospondin 2 sequence and expression
during cell growth and development.";
RL J. Biol. Chem. 267:3274-3281(1992).
RN [2]
RP SEQUENCE OF 1-873 FROM N.A.
RX MEDLINE=91302287; PubMed=1712771;
RA Bornstein P., O'Rourke K., Wikstrom K., Wolf F.W., Katz R., Li P.,
RA Dixit V.M.;
RT "A second, expressed thrombospondin gene (Thbs2) exists in the mouse
genome.";
RL J. Biol. Chem. 266:12821-12824(1991).
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin and type V collagen.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L07803; AAA53064.1; -.
DR EMBL; M64866; AAA40432.1; -.
DR PIR; A42587; A42587.
DR HSSP; P00740; 1EDM.
DR MGD; MGI:98738; Thbs2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSPC.

DR InterPro; IPR003129; TSPN.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00090; tsp_1; 3.
 DR Pfam; PF02412; tsp_3; 13.
 DR Pfam; PF05735; TSPC; 1.
 DR Pfam; PF02210; TSPN; 1.
 DR Pfam; PF00093; vwc; 1.
 DR PRINTS; PRO1705; TSP1REPEAT.
 DR SMART; SM00181; EGF; 3.
 DR SMART; SM00209; TSP1; 3.
 DR SMART; SM00210; TSPN; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS50026; EGF_3; 2.
 DR PROSITE; PS50092; TSP1; 3.
 DR PROSITE; PS01208; VWFC_1; 1.
 DR PROSITE; PS50184; VWFC_2; 1.
 KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
 KW EGF-like domain; Signal.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 1172 THROMBOSPONDIN 2.
 FT DOMAIN 19 215 TSP N-TERMINAL.
 FT DOMAIN 19 232 HEPARIN-BINDING (POTENTIAL).
 FT DOMAIN 318 375 VWFC.
 FT DOMAIN 381 431 TSP TYPE-1 1.
 FT DOMAIN 437 492 TSP TYPE-1 2.
 FT DOMAIN 494 549 TSP TYPE-1 3.
 FT DOMAIN 549 589 EGF-LIKE 1.
 FT DOMAIN 590 647 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 648 692 EGF-LIKE 3.
 FT DOMAIN 725 760 TSP TYPE-3 1.
 FT DOMAIN 761 783 TSP TYPE-3 2.
 FT DOMAIN 784 819 TSP TYPE-3 3.
 FT DOMAIN 820 842 TSP TYPE-3 4.
 FT DOMAIN 843 880 TSP TYPE-3 5.
 FT DOMAIN 881 916 TSP TYPE-3 6.
 FT DOMAIN 917 952 TSP TYPE-3 7.
 FT DOMAIN 953 1172 C-TERMINAL.
 FT SITE 928 930 CELL ATTACHMENT SITE (POTENTIAL).
 FT DISULFID 266 266 INTERCHAIN (PROBABLE).
 FT DISULFID 270 270 INTERCHAIN (PROBABLE).
 FT DISULFID 393 425 BY SIMILARITY.
 FT DISULFID 397 430 BY SIMILARITY.
 FT DISULFID 408 415 BY SIMILARITY.
 FT DISULFID 449 486 BY SIMILARITY.
 FT DISULFID 453 491 BY SIMILARITY.
 FT DISULFID 464 476 BY SIMILARITY.
 FT DISULFID 506 543 BY SIMILARITY.
 FT DISULFID 510 548 BY SIMILARITY.
 FT DISULFID 521 533 BY SIMILARITY.
 FT DISULFID 553 564 BY SIMILARITY.
 FT DISULFID 558 574 BY SIMILARITY.
 FT DISULFID 577 588 BY SIMILARITY.
 FT DISULFID 594 610 BY SIMILARITY.
 FT DISULFID 601 619 BY SIMILARITY.

FT	DISULFID	622	646	BY SIMILARITY.
FT	DISULFID	652	665	BY SIMILARITY.
FT	DISULFID	659	678	BY SIMILARITY.
FT	DISULFID	680	691	BY SIMILARITY.
FT	DISULFID	707	715	BY SIMILARITY.
FT	DISULFID	720	740	BY SIMILARITY.
FT	DISULFID	756	776	BY SIMILARITY.
FT	DISULFID	779	799	BY SIMILARITY.
FT	DISULFID	815	835	BY SIMILARITY.
FT	DISULFID	838	858	BY SIMILARITY.
FT	DISULFID	876	896	BY SIMILARITY.
FT	DISULFID	912	932	BY SIMILARITY.
FT	DISULFID	948	1169	BY SIMILARITY.
FT	CARBOHYD	151	151	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	316	316	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	330	330	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	457	457	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	584	584	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	710	710	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1069	1069	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	1172	AA; 129911 MW;	7CE8E4E8599822AB CRC64;

Query Match 6.1%; Score 293; DB 1; Length 1172;
Best Local Similarity 38.0%; Pred. No. 4.7e-14;
Matches 60; Conservative 22; Mismatches 66; Indels 10; Gaps 5;

Qy	209	RQARLADTANYTCVAKNIVARRRS-ASA AVIVYVN GGWSTWTEWSVCSASCGRGWQKRSR	267
	:: : : : : : :		
Db	403	QRGRSCDVTSNTCLGPSIQTRTC SLGKCDTRIRQNGGWSHWPSSCSVTCGVGNVTRIR	462
Qy	268	SCTNPAPLN GGAFCEGQNVQKTAC-ATLC PVDG SWSPWSKWSACGLDCT---HWRSRECS	323
	: : : : : :		
Db	463	LCNSPVPQMGGKNCKSGRETKPCQRDP CPIDGR WS PWSPWSACTVTCAGGIRERSRVCN	522
Qy	324	DPAPRNGGE ECQG--TD---LDTRNCTS DLCVHSASGP	356
	: : : : : : :		
Db	523	SPEPOYGGKDCVGDVTEHOMCNKRSCP IDGCLS NPCFP	560

RESULT 5

TSP2_BOVIN
ID TSP2_BOVIN STANDARD; PRT; 1170 AA.
AC Q95116; Q28180;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thrombospondin 2 precursor (Corticotropin-induced secreted protein)
(CISP).
GN THBS2 OR TSP2 OR TSP-2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Danik M., Chinn A., Lafeuillade M., Keramidas M., Aquesse-Germon S.,

RA Penhoat A., Chen H., Mosher D., Chambaz E.M., Feige J.J.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE OF 1-522 FROM N.A.
RX MEDLINE=96331130; PubMed=8698834;
RA Lafeuillade B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,
RA Feige J.J.;
RT "Opposite regulation of thrombospondin-1 and corticotropin-induced
RT secreted protein/thrombospondin-2 expression by adrenocorticotrophic
RT hormone in adrenocortical cells.";
RL J. Cell. Physiol. 167:164-172(1996).
RN [3]
RP SEQUENCE OF 318-831 FROM N.A.
RC TISSUE=Aortic endothelium;
RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;
RT "Cloning and sequencing of bovine thrombospondin stimulatory effect of
RT TGF-beta.";
RL Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin and type V collagen.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X96540; CAA65385.1; -.
DR EMBL; X87620; CAA60952.1; -.
DR HSSP; P00740; 1EDM.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSPC.
DR InterPro; IPR003129; TSPN.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; tsp_1; 3.
DR Pfam; PF02412; tsp_3; 13.
DR Pfam; PF05735; TSPC; 1.
DR Pfam; PF02210; TSPN; 1.
DR Pfam; PF00093; vwc; 1.
DR PRINTS; PRO1705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.

DR SMART; SM00209; TSP1; 3.
 DR SMART; SM00210; TSPN; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS50026; EGF_3; 2.
 DR PROSITE; PS50092; TSP1; 3.
 DR PROSITE; PS01208; VWFC_1; 1.
 DR PROSITE; PS50184; VWFC_2; 1.
 KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
 KW EGF-like domain; Signal.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 1170 THROMBOSPONDIN 2.
 FT DOMAIN 19 215 TSP N-TERMINAL.
 FT DOMAIN 19 232 HEPARIN-BINDING (POTENTIAL).
 FT DOMAIN 318 375 VWFC.
 FT DOMAIN 379 429 TSP TYPE-1 1.
 FT DOMAIN 435 490 TSP TYPE-1 2.
 FT DOMAIN 492 547 TSP TYPE-1 3.
 FT DOMAIN 547 587 EGF-LIKE 1.
 FT DOMAIN 588 645 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 646 690 EGF-LIKE 3.
 FT DOMAIN 723 758 TSP TYPE-3 1.
 FT DOMAIN 759 781 TSP TYPE-3 2.
 FT DOMAIN 782 817 TSP TYPE-3 3.
 FT DOMAIN 818 840 TSP TYPE-3 4.
 FT DOMAIN 841 878 TSP TYPE-3 5.
 FT DOMAIN 879 914 TSP TYPE-3 6.
 FT DOMAIN 915 950 TSP TYPE-3 7.
 FT DOMAIN 951 1170 C-TERMINAL.
 FT SITE 926 928 CELL ATTACHMENT SITE (POTENTIAL).
 FT DISULFID 266 266 INTERCHAIN (PROBABLE).
 FT DISULFID 270 270 INTERCHAIN (PROBABLE).
 FT DISULFID 391 423 BY SIMILARITY.
 FT DISULFID 395 428 BY SIMILARITY.
 FT DISULFID 406 413 BY SIMILARITY.
 FT DISULFID 447 484 BY SIMILARITY.
 FT DISULFID 451 489 BY SIMILARITY.
 FT DISULFID 462 474 BY SIMILARITY.
 FT DISULFID 504 541 BY SIMILARITY.
 FT DISULFID 508 546 BY SIMILARITY.
 FT DISULFID 519 531 BY SIMILARITY.
 FT DISULFID 551 562 BY SIMILARITY.
 FT DISULFID 556 572 BY SIMILARITY.
 FT DISULFID 575 586 BY SIMILARITY.
 FT DISULFID 592 608 BY SIMILARITY.
 FT DISULFID 599 617 BY SIMILARITY.
 FT DISULFID 620 644 BY SIMILARITY.
 FT DISULFID 650 663 BY SIMILARITY.
 FT DISULFID 657 676 BY SIMILARITY.
 FT DISULFID 678 689 BY SIMILARITY.
 FT DISULFID 705 713 BY SIMILARITY.
 FT DISULFID 718 738 BY SIMILARITY.
 FT DISULFID 754 774 BY SIMILARITY.
 FT DISULFID 777 797 BY SIMILARITY.
 FT DISULFID 813 833 BY SIMILARITY.
 FT DISULFID 836 856 BY SIMILARITY.

FT DISULFID 874 894 BY SIMILARITY.
 FT DISULFID 910 930 BY SIMILARITY.
 FT DISULFID 946 1167 BY SIMILARITY.
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 316 316 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 455 455 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 582 582 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 708 708 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 936 936 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1067 1067 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 535 535 A -> V (IN REF. 3).
 FT CONFLICT 748 748 S -> T (IN REF. 3).
 SQ SEQUENCE 1170 AA; 129862 MW; 9CF1FBF55B89A051 CRC64;

 Query Match 6.1%; Score 291.5; DB 1; Length 1170;
 Best Local Similarity 38.4%; Pred. No. 6.1e-14;
 Matches 56; Conservative 21; Mismatches 64; Indels 5; Gaps 3;

 Qy 209 RQARLADTANYTCVAKNIVARRSASAA-VIVYVNGGWSTWTEWSVCASCGRGWQKRSR 267
 :: | | : ||: :| | | | : :||| |: || || :|| | | | | | |
 Db 401 QRGRSCDVTSNTCLGSIQTRACSLGRCDHRRIRQDGGSWHSWPSSCSVTCGVGNVTRIR 460

 Qy 268 SCTNPAPLNGGAFCEGQNVQKTAC-ATLCPVDSWSPWSKWSACGLDCT---HWRSRECS 323
 | :| | || |:| : || ||| ||| ||| ||| : | | |:| |:
 Db 461 LCNSPVPQMGGRSCKSGRETAKCQGPPCPVDGRWSPWSACTVTCAGGIRERTRVCN 520

 Qy 324 DPAPRNGGEECQGTLDTRNCTS DLC 349
 | |::| |::| | : : | | |
 Db 521 SPEPQHGGKDCVGGAKEQQMCNRKSC 546

RESULT 6

SM5A_MOUSE
 ID SM5A_MOUSE STANDARD; PRT; 1077 AA.
 AC Q62217;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Semaphorin 5A precursor (Semaphorin F) (Sema F).
 GN SEMA5A OR SEMAF OR SEMF.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NMRI;
 RX MEDLINE=96414430; PubMed=8817451;
 RA Adams R.H., Betz H., Pueschel A.W.;
 RT "A novel class of murine semaphorins with homology to thrombospondin
 RT is differentially expressed during early embryogenesis.";
 RL Mech. Dev. 57:33-45(1996).
 CC -!- FUNCTION: May act as positive axonal guidance cues.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: IN ADULT, DETECTED IN LIVER, BRAIN, KIDNEY,
 HEART, LUNG AND SPLEEN.

CC -!- DEVELOPMENTAL STAGE: DIFFERENTIALLY EXPRESSED IN EMBRYONIC AND
 CC ADULT TISSUES. ITS ABUNDANCE DECREASES FROM E10 TO BITH.
 CC -!- SIMILARITY: Belongs to the semaphorin family.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 7 TSP type-1 domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X97817; CAA66397.1; -.
 DR MGD; MGI:107556; Sema5a.
 DR GO; GO:0016021; C:integral to membrane; IDA.
 DR GO; GO:0008046; F:axon guidance receptor activity; IDA.
 DR GO; GO:0007411; P:axon guidance; IMP.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR002165; Plexin_repeat.
 DR InterPro; IPR001627; Sema.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR Pfam; PF01437; PSI; 1.
 DR Pfam; PF01403; Sema; 1.
 DR Pfam; PF00090; tsp_1; 5.
 DR PRINTS; PRO1705; TSP1REPEAT.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR SMART; SM00209; TSP1; 6.
 DR PROSITE; PS50092; TSP1; 6.
 KW Signal; Transmembrane; Repeat; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 1077 SEMAPHORIN 5A.
 FT DOMAIN 22 971 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 972 992 POTENTIAL.
 FT DOMAIN 993 1077 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 226 507 SEMA.
 FT DOMAIN 540 593 TSP TYPE-1 1.
 FT DOMAIN 595 651 TSP TYPE-1 2.
 FT DOMAIN 653 702 TSP TYPE-1 3.
 FT DOMAIN 707 765 TSP TYPE-1 4.
 FT DOMAIN 784 839 TSP TYPE-1 5.
 FT DOMAIN 841 896 TSP TYPE-1 6.
 FT DOMAIN 897 944 TSP TYPE-1 7.
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 323 323 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 367 367 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 717 717 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 933 933 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 1077 AA; 120826 MW; EDAB0DDDA42789FF CRC64;
 Query Match 6.1%; Score 291; DB 1; Length 1077;
 Best Local Similarity 45.8%; Pred. No. 5.9e-14;
 Matches 54; Conservative 10; Mismatches 50; Indels 4; Gaps 2;

 Qy 241 VNGGWSTWTEWSVCSASCGRGWQKRSRSTNPAPLNGGAFCEGQNVQKTACATL-CPVDG 299
 ||| ||| ||| ||| | || : | | | ||| | || | | ::: | | | |||||
 Db 783 VNGAWSAWTSWSQCSRDCSRGIRNRKRVCNNPEPKFGGMPCLGPSLEFQECNILPCPVDG 842

 Qy 300 SWSPWWSKWSACGLDC---THWRSRECSDPAPRNGGEECQGTDLDRNCTSDLCVHSAS 354
 ||| ||| ||| . | : | : | ||: ||| | |: | | : | : | | |
 Db 843 VWSCWSSWSKCSATCGGGHYMRTSCSNPAPAYGGDICLGLHTEALCNTQTCPESWS 900

RESULT 7
TSP1_XENLA
 ID TSP1_XENLA STANDARD; PRT; 1173 AA.
 AC P35448;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Thrombospondin 1 precursor.
 GN THBS1 OR TSP1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Urry L.A., Ramos J., Duquette M., Desimone D.W., Lawler J.;
 RL Submitted (XXX-1993) to the EMBL/GenBank/DDBJ databases.
 CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
 CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
 CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
 CC V/beta-3 and alpha-IIb/beta-3 (By similarity).
 CC -!- SUBUNIT: Homotrimer; disulfide-linked.
 CC -!- SIMILARITY: Belongs to the thrombospondin family.
 CC -!- SIMILARITY: Contains 1 VWFC domain.
 CC -!- SIMILARITY: Contains 3 EGF-like domains.
 CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
 CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
 CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L04278; -; NOT_ANNOTATED_CDS.
 DR HSSP; P00740; 1EDM.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR006209; EGF_like.

DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR003367; tsp_3.
 DR InterPro; IPR008859; TSPC.
 DR InterPro; IPR003129; TSPN.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00090; tsp_1; 3.
 DR Pfam; PF02412; tsp_3; 13.
 DR Pfam; PF05735; TSPC; 1.
 DR Pfam; PF02210; TSPN; 1.
 DR Pfam; PF00093; vwc; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00209; TSP1; 3.
 DR SMART; SM00210; TSPN; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS50026; EGF_3; 2.
 DR PROSITE; PS50092; TSP1; 3.
 DR PROSITE; PS01208; VWFC_1; 1.
 DR PROSITE; PS50184; VWFC_2; 1.
 KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
 KW EGF-like domain; Signal.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 1173 THROMBOSPONDIN 1.
 FT DOMAIN 23 235 HEPARIN-BINDING (POTENTIAL).
 FT DOMAIN 23 224 TSP N-TERMINAL.
 FT DOMAIN 319 376 VWFC.
 FT DOMAIN 382 432 TSP TYPE-1 1.
 FT DOMAIN 438 493 TSP TYPE-1 2.
 FT DOMAIN 495 550 TSP TYPE-1 3.
 FT DOMAIN 550 590 EGF-LIKE 1.
 FT DOMAIN 591 648 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 649 693 EGF-LIKE 3.
 FT DOMAIN 726 761 TSP TYPE-3 1.
 FT DOMAIN 762 784 TSP TYPE-3 2.
 FT DOMAIN 785 820 TSP TYPE-3 3.
 FT DOMAIN 821 843 TSP TYPE-3 4.
 FT DOMAIN 844 881 TSP TYPE-3 5.
 FT DOMAIN 882 917 TSP TYPE-3 6.
 FT DOMAIN 918 953 TSP TYPE-3 7.
 FT DOMAIN 954 1173 C-TERMINAL.
 FT SITE 929 931 CELL ATTACHMENT SITE (POTENTIAL).
 FT DISULFID 394 426 BY SIMILARITY.
 FT DISULFID 398 431 BY SIMILARITY.
 FT DISULFID 409 416 BY SIMILARITY.
 FT DISULFID 450 487 BY SIMILARITY.
 FT DISULFID 454 492 BY SIMILARITY.
 FT DISULFID 465 477 BY SIMILARITY.
 FT DISULFID 507 544 BY SIMILARITY.
 FT DISULFID 511 549 BY SIMILARITY.
 FT DISULFID 522 534 BY SIMILARITY.
 FT DISULFID 554 565 BY SIMILARITY.
 FT DISULFID 559 575 BY SIMILARITY.

FT	DISULFID	578	589	BY SIMILARITY.
FT	DISULFID	595	611	BY SIMILARITY.
FT	DISULFID	602	620	BY SIMILARITY.
FT	DISULFID	623	647	BY SIMILARITY.
FT	DISULFID	653	666	BY SIMILARITY.
FT	DISULFID	660	679	BY SIMILARITY.
FT	DISULFID	681	692	BY SIMILARITY.
FT	DISULFID	708	716	BY SIMILARITY.
FT	DISULFID	721	741	BY SIMILARITY.
FT	DISULFID	757	777	BY SIMILARITY.
FT	DISULFID	780	800	BY SIMILARITY.
FT	DISULFID	816	836	BY SIMILARITY.
FT	DISULFID	839	859	BY SIMILARITY.
FT	DISULFID	877	897	BY SIMILARITY.
FT	DISULFID	913	933	BY SIMILARITY.
FT	DISULFID	949	1170	BY SIMILARITY.
FT	CARBOHYD	155	155	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	158	158	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	250	250	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	363	363	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	705	705	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	711	711	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1070	1070	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	1173 AA;	130019 MW;	A9F036D6516C0F24 CRC64;

Query Match 6.1%; Score 290; DB 1; Length 1173;
 Best Local Similarity 24.2%; Pred. No. 7.9e-14;
 Matches 92; Conservative 52; Mismatches 144; Indels 92; Gaps 16;

QY	11	LLGIVLAAWLRGSG---AQQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLVCK	66
Db	221	: : : :: :	
QY	221	VFGTTLEAILRNRKGCLSMNTNSVITLDNPVNNGSSPAIRTNIGH-----KTKDLQAVCG	273
QY	67	AVPATQIFFKCNGEWRQVDHVIERSTDGSSGLPTMEVRINVSRRQQVEKVGLEEWYCQC	126
Db	274	: : : : : :	
QY	274	-----FSCDD-----LSKLFAEMKGLRTL---VTTLKDQVTKETEKNELIAQI	313
QY	127	VAWSSSGTTKSQKAYIRIARLRKNFEQ-----EPLAKEVSLEQGIVLPCRP	172
Db	314	: : : : : : :	
QY	314	V-----TRTPGVCLHNGVLHKNRDEWTVDSCTECTCQNSATICRKVSCP---LMPCTN	363
QY	173	---PEG-----IPPAEVEWLRNEDLVDPSDLDPNVYITREHSLVVRQARLADTANYTC	221
Db	364	: : :	
QY	364	ATIPDGECGPCRCWPSDSADDDWSPWSDWTPCS-----VTCGHG-IQQRGRSCDSLNNPC	416
QY	222	VAKNIVAR-----RRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPA	273
Db	417	: : :	
QY	417	EGSSVQTRSCQIQDCDKRFKQ-----DGGWSHWSPWSSCSVTCGSGQITRIRLCNSPV	469
QY	274	PLNGGAFCEGQNQKTA-CATLCPVDGSWPWSKWSACGLDC---THWRSRECSDPAPRN	329
Db	470	:	
QY	470	PQLNGKQCEGEGRENKPCQKDPCPINGQWGPWSLWDTCTVTCGGGMQKRERLCNNPKPQY	529
QY	330	GGEECQGTLDTRNCTS DLC 349	
Db	530	: :	
QY	530	EGKDCIGEPTDSQICNKQDC 549	

RESULT 8

SM5B_HUMAN

ID SM5B_HUMAN STANDARD; PRT; 1093 AA.

AC Q9P283;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Semaphorin 5B precursor.

GN SEMA5B OR KIAA1445.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=20277482; PubMed=10819331;

RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosawa M., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. XVII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.";

RL DNA Res. 7:143-150(2000).

CC --!- FUNCTION: May act as positive axonal guidance cues (By similarity).

CC --!- SUBCELLULAR LOCATION: Type I membrane protein.

CC --!- SIMILARITY: Belongs to the semaphorin family.

CC --!- SIMILARITY: Contains 1 Sema domain.

CC --!- SIMILARITY: Contains 7 TSP type-1 domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AB040878; BAA95969.1; ALT_INIT.

DR Genew; HGNC:10737; SEMA5B.

DR InterPro; IPR003659; Plexin-like.

DR InterPro; IPR002165; Plexin_repeat.

DR InterPro; IPR001627; Sema.

DR InterPro; IPR000884; TSP1.

DR InterPro; IPR008085; TSP_1.

DR Pfam; PF01437; PSI; 1.

DR Pfam; PF01403; Sema; 1.

DR Pfam; PF00090; tsp_1; 5.

DR PRINTS; PRO1705; TSP1REPEAT.

DR SMART; SM00423; PSI; 1.

DR SMART; SM00630; Sema; 1.

DR SMART; SM00209; TSP1; 4.

DR PROSITE; PS50092; TSP1; 5.

KW Signal; Transmembrane; Repeat; Multigene family; Neurogenesis;

KW Developmental protein; Glycoprotein.

FT SIGNAL 1 26 POTENTIAL.

FT CHAIN 27 1093 SEMAPHORIN 5B.

FT DOMAIN 20 978 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 979 999 POTENTIAL.
 FT DOMAIN 1000 1093 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 236 518 SEMA.
 FT DOMAIN 551 605 TSP TYPE-1 1.
 FT DOMAIN 606 662 TSP TYPE-1 2.
 FT DOMAIN 664 713 TSP TYPE-1 3.
 FT DOMAIN 721 776 TSP TYPE-1 4.
 FT DOMAIN 795 850 TSP TYPE-1 5.
 FT DOMAIN 852 907 TSP TYPE-1 6.
 FT DOMAIN 908 952 TSP TYPE-1 7.
 FT CARBOHYD 59 59 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 95 95 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 378 378 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 532 532 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 539 539 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 547 547 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 602 602 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 728 728 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 944 944 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1093 AA; 119866 MW; F1FDEFB87CEAF0EF CRC64;

Query Match 5.8%; Score 276; DB 1; Length 1093;
 Best Local Similarity 31.0%; Pred. No. 8.4e-13;
 Matches 72; Conservative 35; Mismatches 79; Indels 46; Gaps 11;

Qy	241 VNGGWSTWTEWSVCASCGRGWQKRSRSCTNPAPLNGGAFCEGQNQVQKTACAT-LCPVDG 299
	: : : : :
Db	851 VRGAWSCWTWSWSPCSASCAGGHYQRTRRSCTSPAPSPGEDICLGLHTEEALCATQACP--E 908
Qy	300 SWSPWWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDRNCT-SDL-CVHSASGPE 357
	: : : :
Db	909 GWSPWSEWSKCTDDGAQSRSRHCELLP--GSSACAGNSSQSRPCPYSEIPVILPASSME 966
Qy	358 DVALYVG----LIAAVAVCLVL--LLLVLILVYCR--KKEGLDSDVADSSILTSGFQPV 407
	: : : : : : : : :
Db	967 EATGCAGFNLIHLVATGISCFLGSGLLTLAVYLSCQHCQRQSQESTL----- 1013
Qy	408 SIKPSKADNPPLLTIQPDLSTTTTYQGSILCPRQDGPS-P-KFQLTNGHLLSP 458
	: : : : : : : : : :
Db	1014 -VHPATPNHLH-----YKGGGTPKNEKYTPMEFKTLNKNNLIP 1050

RESULT 9

SM5B_MOUSE
 ID SM5B_MOUSE STANDARD; PRT; 1093 AA.
 AC Q60519;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Semaphorin 5B precursor (Semaphorin G) (Sema G).
 GN SEMA5B OR SEMAG OR SEMG.
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NMRI;
 RX MEDLINE=96414430; PubMed=8817451;
 RA Adams R.H., Betz H., Pueschel A.W.;
 RT "A novel class of murine semaphorins with homology to thrombospondin
 is differentially expressed during early embryogenesis.";
 RL Mech. Dev. 57:33-45(1996).
 CC -!-- FUNCTION: May act as positive axonal guidance cues.
 CC -!-- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!-- TISSUE SPECIFICITY: In adult, only detected in brain.
 CC -!-- DEVELOPMENTAL STAGE: Differentially expressed in embryonic and
 CC adult tissues. Its abundance decreases from E10 to birth.
 CC -!-- SIMILARITY: Belongs to the semaphorin family.
 CC -!-- SIMILARITY: Contains 1 Sema domain.
 CC -!-- SIMILARITY: Contains 7 TSP type-1 domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X97818; CAA66398.1; -.
 DR MGD; MGI:107555; Sema5b.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR002165; Plexin_repeat.
 DR InterPro; IPR001627; Sema.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR Pfam; PF01437; PSI; 1.
 DR Pfam; PF01403; Sema; 1.
 DR Pfam; PF00090; tsp_1; 5.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR SMART; SM00209; TSP1; 4.
 DR PROSITE; PS50092; TSP1; 5.
 KW Signal; Transmembrane; Repeat; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 1093 SEMAPHORIN 5B.
 FT DOMAIN 20 978 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 979 999 POTENTIAL.
 FT DOMAIN 1000 1093 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 236 518 SEMA.
 FT DOMAIN 551 605 TSP TYPE-1 1.
 FT DOMAIN 606 662 TSP TYPE-1 2.
 FT DOMAIN 664 713 TSP TYPE-1 3.
 FT DOMAIN 721 776 TSP TYPE-1 4.
 FT DOMAIN 795 850 TSP TYPE-1 5.
 FT DOMAIN 852 907 TSP TYPE-1 6.

FT DOMAIN 908 952 TSP TYPE-1 7.
 FT CARBOHYD 59 59 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 95 95 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 378 378 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 532 532 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 539 539 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 547 547 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 602 602 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 728 728 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 944 944 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1093 AA; 120326 MW; 29E5C9B1E8108717 CRC64;

 Query Match 5.8%; Score 275.5; DB 1; Length 1093;
 Best Local Similarity 32.1%; Pred. No. 9.2e-13;
 Matches 69; Conservative 18; Mismatches 75; Indels 53; Gaps 8;

 Qy 163 EQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLA
 Db 737 EQRFRFTCRAP-----LPDP-----HGLQFGKRR--TETRTCP 767

 Qy 223 AKNIVA-----RRRSASA
 Db 768 ADGTGACDTDALVEDLLRSGSTSPHTL--NGGWATWGPWSSCSRDCELGFRVRKRTCTN 824

 Qy 272 PAPLN
 Db 825 PEPRNGGLPCVGDAAEYQDCNPQACPVRGAWSQCSASC

 Qy 328 RNGGEECQGT
 Db 885 SPGEDICLGLHTEEALCSTQAC-----PEGWSLW 913

RESULT 10
BAI3_HUMAN
 ID BAI3_HUMAN STANDARD; PRT; 1522 AA.
 AC O60242; O60297;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Brain-specific angiogenesis inhibitor 3 precursor.
 GN BAI3 OR KIAA0550.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=98194217; PubMed=9533023;
 RA Shiratsuchi T., Nishimori H., Ichise H., Nakamura Y., Tokino T.;
 RT "Cloning and characterization of BAI2 and BAI3, novel genes homologous
 to brain-specific angiogenesis inhibitor 1 (BAI1).";

RL Cytogenet. Cell Genet. 79:103-108(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98290545; PubMed=9628581;
RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. IX.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 5:31-39(1998).
RN [3]
RP SEQUENCE FROM N.A., AND REVISIONS TO 643-665 AND C-TERMINUS.
RX MEDLINE=22158633; PubMed=12168954;
RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
RT "Construction of expression-ready cDNA clones for KIAA genes: manual
RT curation of 330 KIAA cDNA clones.";
RL DNA Res. 9:99-106(2002).
CC -!- FUNCTION: MIGHT BE INVOLVED IN ANGIOGENESIS INHIBITION AND
CC SUPPRESSION OF GLIOBLASTOMA.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN BRAIN. ALSO DETECTED IN
CC HEART. REDUCED EXPRESSION IS OBSERVED IN SOME GLIOBLASTOMA CELL
CC LINES.
CC -!- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 4 TSP type-1 domains.
CC -!- SIMILARITY: Contains 1 GPS domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB005299; BAA25363.1; -.
DR EMBL; AB011122; BAA25476.2; ALT_INIT.
DR PIR; T00028; T00028.
DR Genew; HGNC:945; BAI3.
DR MIM; 602684; -.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR000832; GPCR_secretin.
DR InterPro; IPR001879; hormn_receptor.
DR InterPro; IPR000203; PKD_cys_rich.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00002; 7tm_2; 1.
DR Pfam; PF01825; GPS; 1.
DR Pfam; PF02793; HRM; 1.
DR Pfam; PF00090; tsp_1; 4.
DR SMART; SM00303; GPS; 1.
DR SMART; SM00008; HormR; 1.
DR SMART; SM00209; TSP1; 4.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS50221; GPS; 1.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; FALSE_NEG.

DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; FALSE_NEG.
 DR PROSITE; PS50227; G_PROTEIN_RECEP_F2_3; 1.
 DR PROSITE; PS50261; G_PROTEIN_RECEP_F2_4; 1.
 DR PROSITE; PS50092; TSP1; 4.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Repeat.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 1522 BRAIN-SPECIFIC ANGIOGENESIS INHIBITOR 3.
 FT DOMAIN 25 880 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 881 901 1 (POTENTIAL).
 FT DOMAIN 902 910 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 911 931 2 (POTENTIAL).
 FT DOMAIN 932 939 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 940 960 3 (POTENTIAL).
 FT DOMAIN 961 981 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 982 1002 4 (POTENTIAL).
 FT DOMAIN 1003 1023 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1024 1044 5 (POTENTIAL).
 FT DOMAIN 1045 1098 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 1099 1119 6 (POTENTIAL).
 FT DOMAIN 1120 1125 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1126 1146 7 (POTENTIAL).
 FT DOMAIN 1147 1522 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 30 159 CUB.
 FT DOMAIN 291 343 TSP TYPE-1 1.
 FT DOMAIN 345 398 TSP TYPE-1 2.
 FT DOMAIN 400 453 TSP TYPE-1 3.
 FT DOMAIN 455 508 TSP TYPE-1 4.
 FT DOMAIN 816 868 GPS.
 FT DOMAIN 942 945 POLY-THR.
 FT DOMAIN 1173 1176 POLY-SER.
 FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 337 337 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 540 540 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 779 779 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 828 828 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 937 937 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1522 AA; 171490 MW; D22D0A5D4BB62502 CRC64;

Query Match 5.7%; Score 275; DB 1; Length 1522;
 Best Local Similarity 39.0%; Pred. No. 1.6e-12;
 Matches 57; Conservative 20; Mismatches 53; Indels 16; Gaps 6;

Qy 220 TCVA-----KNIVARRRSASA
AVIVYVNGGWSTWTEWSVC
SASCGRGWQKRSRSCTNPA 273
|||:/ : | : : ||: | | : ||:|| :|||| : ||:|||
 Db 317 TCVSPYGTHCSGPLRESRVCNN
TALCPVHG
VWE
WSPWSLCSFTCGRGQRT
RTRSCT--P 374

Qy 274 PLNGGAFCEGQN
VQKTAC-ATLC
PVDGSWSPWS
KWSACGLDC--THW
RSRECSDPAPRN 329
| || ||| | ||| ||| | || || | : | | ||| :| : | :
 Db 375 PQYGGRPCEGPETHHKPCNIALCP
VDGQWQEWS
SQCSVTC
SNGTQQRSRQCT--AAA
AH 432

QY 330 GGEECQGTLDTRNCTSDLCVHSASG 355
|| ||:| ::| | : | :|:
Db 433 GGSECRGPWAESRECYNPEC--TANG 456

RESULT 11
BAI2_HUMAN
ID BAI2_HUMAN STANDARD; PRT; 1572 AA.
AC O60241;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Brain-specific angiogenesis inhibitor 2 precursor.
GN BAI2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=98194217; PubMed=9533023;
RA Shiratsuchi T., Nishimori H., Ichise H., Nakamura Y., Tokino T.;
RT "Cloning and characterization of BAI2 and BAI3, novel genes homologous
to brain-specific angiogenesis inhibitor 1 (BAI1).";
RL Cytogenet. Cell Genet. 79:103-108(1997).
CC --!- FUNCTION: MIGHT BE INVOLVED IN ANGIOGENESIS INHIBITION.
CC --!- SUBCELLULAR LOCATION: Integral membrane protein.
CC --!- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN BRAIN. ALSO DETECTED IN
CC HEART, THYMUS, SKELETAL MUSCLE, AND DIFFERENT CELL LINES.
CC --!- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.
CC --!- SIMILARITY: Contains 4 TSP type-1 domains.
CC --!- SIMILARITY: Contains 1 GPS domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB005298; BAA25362.1; -.
DR PIR; T00027; T00027.
DR Genew; HGNC:944; BAI2.
DR MIM; 602683; -.
DR InterPro; IPR000832; GPCR_secretin.
DR InterPro; IPR001879; hormn_receptor.
DR InterPro; IPR000203; PKD_cys_rich.
DR InterPro; IPR000884; TSPI.
DR Pfam; PF00002; 7tm_2; 1.
DR Pfam; PF01825; GPS; 1.
DR Pfam; PF02793; HRM; 1.
DR Pfam; PF00090; tsp_1; 4.
DR SMART; SM00303; GPS; 1.
DR SMART; SM00008; HormR; 1.

Db 324 CVSSPYGTLCSGPLRETRPCNNSATCPVHGVEEWGSWSLCSRSGRSRMRVCV--P 381
 Qy 274 PLNGGAFCEGQNVQKTACA-TLCPVDGWSPWSKWSACGLDC---THWRSRECSDPAPR- 328
 | :|| ||| :| |: | ||:| | | | | | | |||:|| |
 Db 382 PQHGGKACEGPELQTKLCSMAACPVEGQWLEWGPWGPCSTSCANGTQQRSRKCSVAGPAW 441
 Qy 329 -----NGGEECQ 335
 | |:
 Db 442 ATCTGALTDTRECSNLECPATDSKWGPWNAWSLCSKTCGTGWQRRFRMCQATGTQGYPCE 501
 Qy 336 GTDLDTRNCTS DLC--VHSASGPEDVAL----- 361
 | | : : |: | | | | | |
 Db 502 GTGEEVKPCSEKRCPAFHEMCRDEYVMLMTWKAAAGEIIYNKCPPNASGSASRRCLLSA 561
 Qy 362 ----YVGLIAVAVCL--VLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKA 414
 | | | : | |: | | : : | : : | | :: | | : : :
 Db 562 QGVAYWGLPSFARCISHEYRYLYLSREHLAKGQRMLAGEGMSQVVRSLQELLARRYY 620
 Qy 415 DNPHLLTIQPDLS TTTTYQGS LCP RQDG PSpKFQLT-----NGHLLSP LGG 461
 | :: : | | : : | | | | | | :| : | | :| |
 Db 621 SGDLLFSVDILRNVTDTFKRATYVPSADDVQRFFQVVS FMVDAENKEWDDAQQVSP--G 678
 Qy 462 GRHTLHHSSPTSEAEVF-----SRLSTQNYFRSLPRG----TSNMTYGT FN 504
 | | | :| :| | : | | | | | | :| :| :| :|:
 Db 679 SVHLLR-----VVEDFIHLVGDALKAFQSSLIVTDNLVISIQREP VSAVSSDITFP MRG 732
 Qy 505 FLG-----GRLMI PNTG I SLLIP-----PDAI PRG K----- 530
 | | | :| :| | | | | | | | | | | | :| |
 Db 733 RRGMKD WVRH SEDR LFLPKEV I SLS PGK PATSGAAGG PGRGRGPGTV PPG GHSH QRLL 792
 Qy 531 -----IYE-IYLT LH KPEDV RLPLAG C QT LLSP IV SC GPPG VLL TRPV IL 574
 | : | | | | | | | | | | | | | | | | | |:
 Db 793 PADPDESSYF VIG A VLYRTL GL LIP PPP--RPPLAV TS RV MT--VTV RP PT Q PPA EPLIT 847
 Qy 575 A-----MDHC GE P SP DS WS LRL KKQS CEG SW ED VL HL GEE AP SH LY YCQ- LE AS AC YV-- 626
 | : | | | | | | | | | | | | | | | | | | |:
 Db 848 VEL SYI ING TT DPHC AS WDYS-RADASS GDWD-----TENC Q LET Q AA HTRC 894
 Qy 627 FTEQLGRF ALVGE----ALS VAA KRL KLL FAP VACTSLEY NIRV YCLHD TH DAL KEV 681
 | : | | :| :| | :| :| :| :| :| :| :| |
 Db 895 QCQHLSTFAVLAQPPKDLTLELAGSPSVPLVIGCAVSCM ALLT LAI YA-----AF WRF 948
 Qy 682 VQLEKQLGGQLI QE P RVLH FKDSYH NLRLSI HDV PSS LWKS KLL VS YQE I P FYHI WNGT Q 741
 | : | : | | | | | | | | | | | | | | | | | |:
 Db 949 IKSERSI-----ILLNF CLSI--LAS NI--LILVGQSRVLSKG VCT MTA 988
 Qy 742 RYLHCTFTLERVSPSTS DLACKLWV-----WQVEGDG 773
 | :| | | | | | | | | | | | | | | | | | | |:
 Db 989 AFLHFFF-----LSSFCW VL TEAW QSYL AVI GRM RTRL VRK RFL CLG WGL P ALV 1037
 Qy 774 QSFSINFN ITKDTRFA ELLALESEAG-VPALVGPSA-----FKIPFLIRQKI----IS 821
 | : | | | | | | | | | | | | | | | | | | | | |:
 Db 1038 VAVSVGFTRTKG YGTSSY CWL SLEG GLL YAFVGPA AVI VL VNMLIGI IV FN KLM ARDG IS 1097
 Qy 822 SLDPPCRRGAD---WRT L 836
 | | :| | :|
 Db 1098 DKS KKQ RAGSER CPW ASL 1115

RESULT 12

TSP1_MOUSE

ID TSP1_MOUSE STANDARD; PRT; 1170 AA.
AC P35441;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thrombospondin 1 precursor.
GN THBS1 OR TSP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92128941; PubMed=1774063;
RA Lawler J., Duquette M., Ferro P., Copeland N.G., Gilbert D.J.,
RA Jenkins N.A.;
RT "Characterization of the murine thrombospondin gene.";
RL Genomics 11:587-600(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92147683; PubMed=1371115;
RA Laherty C.D., O'Rourke K., Wolf F.W., Katz R., Seldin M.F.,
RA Dixit V.M.;
RT "Characterization of mouse thrombospondin 2 sequence and expression
during cell growth and development.";
RL J. Biol. Chem. 267:3274-3281(1992).
RN [3]
RP SEQUENCE OF 1-490 FROM N.A.
RX MEDLINE=90375546; PubMed=2398070;
RA Bornstein P., Alfi D., Devarayalu S., Framson P., Li P.;
RT "Characterization of the mouse thrombospondin gene and evaluation of
the role of the first intron in human gene expression.";
RL J. Biol. Chem. 265:16691-16698(1990).
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-IIb/beta-3.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M62470; AAA50611.1; -.
DR EMBL; M62450; AAA50611.1; JOINED.
DR EMBL; M62451; AAA50611.1; JOINED.
DR EMBL; M62452; AAA50611.1; JOINED.
DR EMBL; M62453; AAA50611.1; JOINED.
DR EMBL; M62454; AAA50611.1; JOINED.
DR EMBL; M62455; AAA50611.1; JOINED.
DR EMBL; M62456; AAA50611.1; JOINED.
DR EMBL; M62457; AAA50611.1; JOINED.
DR EMBL; M62458; AAA50611.1; JOINED.
DR EMBL; M62459; AAA50611.1; JOINED.
DR EMBL; M62460; AAA50611.1; JOINED.
DR EMBL; M62461; AAA50611.1; JOINED.
DR EMBL; M62462; AAA50611.1; JOINED.
DR EMBL; M62463; AAA50611.1; JOINED.
DR EMBL; M62464; AAA50611.1; JOINED.
DR EMBL; M62465; AAA50611.1; JOINED.
DR EMBL; M62466; AAA50611.1; JOINED.
DR EMBL; M62467; AAA50611.1; JOINED.
DR EMBL; M62468; AAA50611.1; JOINED.
DR EMBL; M62469; AAA50611.1; JOINED.
DR EMBL; M87276; AAA53063.1; -.
DR EMBL; J05606; AAA40431.1; -.
DR EMBL; J05605; AAA40431.1; JOINED.
DR PIR; A40558; A40558.
DR MGD; MGI:98737; Thbs1.
DR GO; GO:0005615; C:extracellular space; IDA.
DR GO; GO:0016525; P:negative regulation of angiogenesis; IDA.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSPC.
DR InterPro; IPR003129; TSPN.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; tsp_1; 3.
DR Pfam; PF02412; tsp_3; 13.
DR Pfam; PF05735; TSPC; 1.
DR Pfam; PF02210; TSPN; 1.
DR Pfam; PF00093; vwc; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS50184; VWFC_2; 1.
KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW EGF-like domain; Signal.
FT SIGNAL 1 18 POTENTIAL.

FT	CHAIN	19	1170	THROMBOSPONDIN 1.
FT	DOMAIN	19	232	HEPARIN-BINDING (POTENTIAL).
FT	DOMAIN	24	221	TSP N-TERMINAL.
FT	DOMAIN	316	373	VWFC.
FT	DOMAIN	379	429	TSP TYPE-1 1.
FT	DOMAIN	435	490	TSP TYPE-1 2.
FT	DOMAIN	492	547	TSP TYPE-1 3.
FT	DOMAIN	549	587	EGF-LIKE 1.
FT	DOMAIN	588	645	EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	646	690	EGF-LIKE 3.
FT	DOMAIN	723	758	TSP TYPE-3 1.
FT	DOMAIN	759	781	TSP TYPE-3 2.
FT	DOMAIN	782	817	TSP TYPE-3 3.
FT	DOMAIN	818	840	TSP TYPE-3 4.
FT	DOMAIN	841	878	TSP TYPE-3 5.
FT	DOMAIN	879	914	TSP TYPE-3 6.
FT	DOMAIN	915	950	TSP TYPE-3 7.
FT	DOMAIN	951	1170	C-TERMINAL.
FT	SITE	926	928	CELL ATTACHMENT SITE (POTENTIAL).
FT	DISULFID	270	270	INTERCHAIN (PROBABLE).
FT	DISULFID	274	274	INTERCHAIN (PROBABLE).
FT	DISULFID	391	423	BY SIMILARITY.
FT	DISULFID	395	428	BY SIMILARITY.
FT	DISULFID	406	413	BY SIMILARITY.
FT	DISULFID	447	484	BY SIMILARITY.
FT	DISULFID	451	489	BY SIMILARITY.
FT	DISULFID	462	474	BY SIMILARITY.
FT	DISULFID	504	541	BY SIMILARITY.
FT	DISULFID	508	546	BY SIMILARITY.
FT	DISULFID	519	531	BY SIMILARITY.
FT	DISULFID	551	562	BY SIMILARITY.
FT	DISULFID	556	572	BY SIMILARITY.
FT	DISULFID	575	586	BY SIMILARITY.
FT	DISULFID	592	608	BY SIMILARITY.
FT	DISULFID	599	617	BY SIMILARITY.
FT	DISULFID	620	644	BY SIMILARITY.
FT	DISULFID	650	663	BY SIMILARITY.
FT	DISULFID	657	676	BY SIMILARITY.
FT	DISULFID	678	689	BY SIMILARITY.
FT	DISULFID	705	713	BY SIMILARITY.
FT	DISULFID	718	738	BY SIMILARITY.
FT	DISULFID	754	774	BY SIMILARITY.
FT	DISULFID	777	797	BY SIMILARITY.
FT	DISULFID	813	833	BY SIMILARITY.
FT	DISULFID	836	856	BY SIMILARITY.
FT	DISULFID	874	894	BY SIMILARITY.
FT	DISULFID	910	930	BY SIMILARITY.
FT	DISULFID	946	1167	BY SIMILARITY.
FT	CARBOHYD	248	248	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	360	360	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	708	708	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1067	1067	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	1025	1025	F -> L (IN REF. 2).
SQ	SEQUENCE	1170 AA;	129646 MW;	0443E493615E7F06 CRC64;

Query Match 5.6%; Score 270.5; DB 1; Length 1170;
 Best Local Similarity 32.2%; Pred. No. 2.4e-12;

Matches 57; Conservative 24; Mismatches 71; Indels 25; Gaps 5;
 Qy 207 VVRQARLADTANYTCVAKNIVAR-----RRSASAIVYVNGGWSTWEWSVCASC 258
 : :: | |: | | :: | :| :||| |: || || :|
 Db 399 IQQQRGRSCDSLNNRCEGSSVQTRTCHIQECDKRFKQ-----DGGWSHWSPWSSCSVTC 451
 Qy 259 GRGWQKRSRSCTNPAPLNGGAFCEGQNVQKTAC-ATLCPVDGSWSPWSKWSACGLDC--- 314
 | | | | | :|: | | ||: : | || |::| | ||| | | : |
 Db 452 GDGVITRIRLNCNSPSPQMNGKPCEGEARETKACKDACPINGGWGPWSPWDICSVTCGGG 511
 Qy 315 THWRSRECSDPAPRNGGEECQGTDLDTRNCTS DLCVHSASGPEDVALYVGLIAVAVC 371
 ||| |::| |: ||::| | : : | | | | | | | | | | |
 Db 512 VQRRSRLCNNPTPQFGGKDCVGDVTENQVCNKQDC-----PIDGCLSNPCFAGAKC 562

RESULT 13

TSP1_HUMAN

ID TSP1_HUMAN STANDARD; PRT; 1170 AA.
 AC P07996; Q15667;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Thrombospondin 1 precursor.
 GN THBS1 OR TSP1 OR TSP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Endothelial cells;
 RX MEDLINE=87057617; PubMed=2430973;
 RA Lawler J., Hynes R.O.;
 RT "The structure of human thrombospondin, an adhesive glycoprotein with
 multiple calcium-binding sites and homologies with several different
 proteins.";
 RL J. Cell Biol. 103:1635-1648(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89139590; PubMed=2918029;
 RA Hennessy S.W., Frazier B.A., Kim D.D., Deckwerth T.L.,
 RA Baumgartel D.M., Rotwein P., Frazier W.A.;
 RT "Complete thrombospondin mRNA sequence includes potential regulatory
 sites in the 3' untranslated region.";
 RL J. Cell Biol. 108:729-736(1989).
 RN [3]
 RP SEQUENCE OF 1-397 FROM N.A.
 RX MEDLINE=87157592; PubMed=3030396;
 RA Kobayashi S., Eden-Mccutchan F., Framson P., Bornstein P.;
 RT "Partial amino acid sequence of human thrombospondin as determined by
 analysis of cDNA clones: homology to malarial circumsporozoite
 proteins.";
 RL Biochemistry 25:8418-8425(1986).
 RN [4]
 RP SEQUENCE OF 1-374 FROM N.A.
 RX MEDLINE=86287276; PubMed=3461443;
 RA Dixit V.M., Hennessy S.W., Grant G.A., Rotwein P., Frazier W.A. ;

RT "Characterization of a cDNA encoding the heparin and collagen binding
RT domains of human thrombospondin.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5449-5453(1986).
RN [5]
RP SEQUENCE OF 1-166 FROM N.A.
RX MEDLINE=89291870; PubMed=2544587;
RA Laherty C.D., Gierman T.M., Dixit V.M.;
RT "Characterization of the promoter region of the human thrombospondin
RT gene. DNA sequences within the first intron increase transcription.";
RL J. Biol. Chem. 264:11222-11227(1989).
RN [6]
RP SEQUENCE OF 1028-1170 FROM N.A.
RA la Fleur M., Jobin C., Gauthier J., Kreis C.G.;
RL Submitted (XXX-1992) to the EMBL/GenBank/DDBJ databases.
RN [7]
RP CARBOHYDRATE-LINKAGE SITES TRP-385; SER-394; TRP-438; TRP-441;
RP THR-450; TRP-498 AND THR-507.
RC TISSUE=Platelet;
RX MEDLINE=21125860; PubMed=11067851;
RA Hofsteenge J., Huwiler K.G., Macek B., Hess D., Lawler J.,
RA Mosher D.F., Peter-Katalinic J.;
RT "C-mannosylation and O-fucosylation of the thrombospondin type 1
RT module.";
RL J. Biol. Chem. 276:6485-6498(2001).
RN [8]
RP THROMBOSPONDIN DOMAIN DISULFIDE BRIDGES.
RX MEDLINE=22338361; PubMed=12450399;
RA Huwiler K.G., Vestling M.M., Annis D.S., Mosher D.F.;
RT "Biophysical characterization, including disulfide bond assignments,
RT of the anti-angiogenic type 1 domains of human thrombospondin-1.";
RL Biochemistry 41:14329-14339(2002).
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-IIb/beta-3.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M25631; AAA36741.1; -.
DR EMBL; X04665; CAA28370.1; -.
DR EMBL; X14787; CAA32889.1; -.
DR EMBL; M14326; AAA61237.1; ALT_SEQ.
DR EMBL; J04835; AAA61178.1; -.
DR EMBL; M99425; AAB59366.1; -.

DR PIR; A26155; TSHUP1.
 DR PDB; 1LSL; 18-DEC-02.
 DR GlycoSuiteDB; P07996; -.
 DR Genew; HGNC:11785; THBS1.
 DR MIM; 188060; -.
 DR GO; GO:0004866; F:endopeptidase inhibitor activity; TAS.
 DR GO; GO:0004871; F:signal transducer activity; TAS.
 DR GO; GO:0007275; P:development; TAS.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR003367; tsp_3.
 DR InterPro; IPR008859; TSPC.
 DR InterPro; IPR003129; TSPN.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00090; tsp_1; 3.
 DR Pfam; PF02412; tsp_3; 13.
 DR Pfam; PF05735; TSPC; 1.
 DR Pfam; PF02210; TSPN; 1.
 DR Pfam; PF00093; vwc; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00181; EGF; 3.
 DR SMART; SM00209; TSP1; 3.
 DR SMART; SM00210; TSPN; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS50026; EGF_3; 2.
 DR PROSITE; PS50092; TSP1; 3.
 DR PROSITE; PS01208; VWFC_1; 1.
 DR PROSITE; PS50184; VWFC_2; 1.
 KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
 KW EGF-like domain; Signal; 3D-structure.
 FT SIGNAL 1 18
 FT CHAIN 19 1170 THROMBOSPONDIN 1.
 FT DOMAIN 19 232 HEPARIN-BINDING (POTENTIAL).
 FT DOMAIN 24 221 TSP N-TERMINAL.
 FT DOMAIN 316 373 VWFC.
 FT DOMAIN 379 429 TSP TYPE-1 1.
 FT DOMAIN 435 490 TSP TYPE-1 2.
 FT DOMAIN 492 547 TSP TYPE-1 3.
 FT DOMAIN 549 587 EGF-LIKE 1.
 FT DOMAIN 588 645 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 646 690 EGF-LIKE 3.
 FT DOMAIN 723 758 TSP TYPE-3 1.
 FT DOMAIN 759 781 TSP TYPE-3 2.
 FT DOMAIN 782 817 TSP TYPE-3 3.
 FT DOMAIN 818 840 TSP TYPE-3 4.
 FT DOMAIN 841 878 TSP TYPE-3 5.
 FT DOMAIN 879 914 TSP TYPE-3 6.
 FT DOMAIN 915 950 TSP TYPE-3 7.
 FT DOMAIN 951 1170 C-TERMINAL.
 FT SITE 926 928 CELL ATTACHMENT SITE (POTENTIAL).
 FT DISULFID 270 270 INTERCHAIN (PROBABLE).

FT	DISULFID	274	274	INTERCHAIN (PROBABLE).
FT	DISULFID	391	423	
FT	DISULFID	395	428	
FT	DISULFID	406	413	
FT	DISULFID	447	484	
FT	DISULFID	451	489	
FT	DISULFID	462	474	
FT	DISULFID	504	541	
FT	DISULFID	508	546	
FT	DISULFID	519	531	
FT	DISULFID	551	562	BY SIMILARITY.
FT	DISULFID	556	572	BY SIMILARITY.
FT	DISULFID	575	586	BY SIMILARITY.
FT	DISULFID	592	608	BY SIMILARITY.
FT	DISULFID	599	617	BY SIMILARITY.
FT	DISULFID	620	644	BY SIMILARITY.
FT	DISULFID	650	663	BY SIMILARITY.
FT	DISULFID	657	676	BY SIMILARITY.
FT	DISULFID	678	689	BY SIMILARITY.
FT	DISULFID	705	713	BY SIMILARITY.
FT	DISULFID	718	738	BY SIMILARITY.
FT	DISULFID	754	774	BY SIMILARITY.
FT	DISULFID	777	797	BY SIMILARITY.
FT	DISULFID	813	833	BY SIMILARITY.
FT	DISULFID	836	856	BY SIMILARITY.
FT	DISULFID	874	894	BY SIMILARITY.
FT	DISULFID	910	930	BY SIMILARITY.
FT	DISULFID	946	1167	BY SIMILARITY.
FT	CARBOHYD	248	248	N-LINKED (GLCNAC. . .).
FT	CARBOHYD	360	360	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	385	385	C-LINKED (MAN).
FT				/FTId=CAR_000205.
FT	CARBOHYD	394	394	O-LINKED (FUC. . .).
FT				/FTId=CAR_000206.
FT	CARBOHYD	438	438	C-LINKED (MAN).
FT				/FTId=CAR_000207.
FT	CARBOHYD	441	441	C-LINKED (MAN).
FT				/FTId=CAR_000208.
FT	CARBOHYD	450	450	O-LINKED (FUC. . .).
FT				/FTId=CAR_000209.
FT	CARBOHYD	498	498	C-LINKED (MAN).
FT				/FTId=CAR_000210.
FT	CARBOHYD	507	507	O-LINKED (FUC. . .).
FT				/FTId=CAR_000211.
FT	CARBOHYD	708	708	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1067	1067	N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 5.6%; Score 268.5; DB 1; Length 1170;
 Best Local Similarity 32.9%; Pred. No. 3.4e-12;
 Matches 51; Conservative 24; Mismatches 61; Indels 19; Gaps 4;

Qy 207 VVRQARLADTANYTCVAKNIVAR-----RRSASAIVIVYNGGWSTWESVCSASC 258
 : :: | | : | | :: | | : | : | | | | | : |

Db 399 IQQQRGRSCDSLNNRCEGSSVQTRTCHIQECDKRFKQ-----DGGWSHWSPWSSCSVTC 451

Qy 259 GRGWQKRRSRCTNPAPLNGGAFCEGQNVQKTAC-ATLCPVDGWSWSKWSACGLDC--- 314
 | | | | | : | : | | | | : | : | | | | | | | : |

Db 452 GDGVITRIRLCNSPSQMNGKPCEGEARETKACKDACPINGGWGPWSPWDICSVTCGGG 511
Qy 315 THWRSRECSDPAPRNGGEECQGTLDTRNCTS DLC 349
||| ::| |: ||::| | : : | |
Db 512 VQKRSRLCNNPTPQFGGKDCVGDVTENQICNKQDC 546

RESULT 14
TSP1_BOVIN
ID TSP1_BOVIN STANDARD; PRT; 1170 AA.
AC Q28178; Q28179;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Thrombospondin 1 precursor.
GN THBS1 OR TSP1 OR TSP-1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Holstein; TISSUE=Tooth;
RX MEDLINE=98173773; PubMed=9507054;
RA Ueno A., Yamashita K., Nagata T., Tsurumi C., Miwa Y., Kitamura S.,
RA Inoue H.;
RT "cDNA cloning of bovine thrombospondin 1 and its expression in
RT odontoblasts and predentin.";
RL Biochim. Biophys. Acta 1382:17-22(1998).
RN [2]
RP SEQUENCE OF 1-18 AND 710-1170 FROM N.A.
RC TISSUE=Aortic endothelium;
RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-IIb/beta-3. May play a role in dentinogenesis
CC and/or maintenance of dentin and dental pulp.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- TISSUE SPECIFICITY: Odontoblasts.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AB005287; BAA21115.1; -.
 DR EMBL; X87618; CAA60950.1; -.
 DR EMBL; X87619; CAA60951.1; -.
 DR PIR; S55501; S55501.
 DR GlycoSuitedB; Q28178; -.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR InterPro; IPR003367; tsp_3.
 DR InterPro; IPR008859; TSPC.
 DR InterPro; IPR003129; TSPN.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00090; tsp_1; 3.
 DR Pfam; PF02412; tsp_3; 13.
 DR Pfam; PF05735; TSPC; 1.
 DR Pfam; PF02210; TSPN; 1.
 DR Pfam; PF00093; vwc; 1.
 DR PRINTS; PRO1705; TSP1REPEAT.
 DR SMART; SM00181; EGF; 3.
 DR SMART; SM00209; TSP1; 3.
 DR SMART; SM00210; TSPN; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS50026; EGF_3; 2.
 DR PROSITE; PS50092; TSP1; 3.
 DR PROSITE; PS01208; VWFC_1; 1.
 DR PROSITE; PS50184; VWFC_2; 1.
 KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
 KW EGF-like domain; Signal.
 FT SIGNAL 1 18 BY SIMILARITY.
 FT CHAIN 19 1170 THROMBOSPONDIN 1.
 FT DOMAIN 19 232 HEPARIN-BINDING (POTENTIAL).
 FT DOMAIN 24 221 TSP N-TERMINAL.
 FT DOMAIN 316 373 VWFC.
 FT DOMAIN 379 429 TSP TYPE-1 1.
 FT DOMAIN 435 490 TSP TYPE-1 2.
 FT DOMAIN 492 547 TSP TYPE-1 3.
 FT DOMAIN 549 587 EGF-LIKE 1.
 FT DOMAIN 588 645 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 646 690 EGF-LIKE 3.
 FT DOMAIN 723 758 TSP TYPE-3 1.
 FT DOMAIN 759 781 TSP TYPE-3 2.
 FT DOMAIN 782 817 TSP TYPE-3 3.
 FT DOMAIN 818 840 TSP TYPE-3 4.
 FT DOMAIN 841 878 TSP TYPE-3 5.
 FT DOMAIN 879 914 TSP TYPE-3 6.
 FT DOMAIN 915 950 TSP TYPE-3 7.
 FT DOMAIN 951 1170 C-TERMINAL.
 FT SITE 926 928 CELL ATTACHMENT SITE (POTENTIAL).
 FT DISULFID 270 270 INTERCHAIN (PROBABLE).
 FT DISULFID 274 274 INTERCHAIN (PROBABLE).
 FT DISULFID 391 423 BY SIMILARITY.
 FT DISULFID 395 428 BY SIMILARITY.

FT	DISULFID	406	413	BY SIMILARITY.
FT	DISULFID	447	484	BY SIMILARITY.
FT	DISULFID	451	489	BY SIMILARITY.
FT	DISULFID	462	474	BY SIMILARITY.
FT	DISULFID	504	541	BY SIMILARITY.
FT	DISULFID	508	546	BY SIMILARITY.
FT	DISULFID	519	531	BY SIMILARITY.
FT	DISULFID	551	562	BY SIMILARITY.
FT	DISULFID	556	572	BY SIMILARITY.
FT	DISULFID	575	586	BY SIMILARITY.
FT	DISULFID	592	608	BY SIMILARITY.
FT	DISULFID	599	617	BY SIMILARITY.
FT	DISULFID	620	644	BY SIMILARITY.
FT	DISULFID	650	663	BY SIMILARITY.
FT	DISULFID	657	676	BY SIMILARITY.
FT	DISULFID	678	689	BY SIMILARITY.
FT	DISULFID	705	713	BY SIMILARITY.
FT	DISULFID	718	738	BY SIMILARITY.
FT	DISULFID	754	774	BY SIMILARITY.
FT	DISULFID	777	797	BY SIMILARITY.
FT	DISULFID	813	833	BY SIMILARITY.
FT	DISULFID	836	856	BY SIMILARITY.
FT	DISULFID	874	894	BY SIMILARITY.
FT	DISULFID	910	930	BY SIMILARITY.
FT	DISULFID	946	1167	BY SIMILARITY.
FT	CARBOHYD	248	248	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	360	360	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	708	708	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1067	1067	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1085	1085	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	805	805	S -> G (IN REF. 2).
SQ	SEQUENCE	1170 AA;	129533 MW;	ODD6ADF3E5FA031A CRC64;

Query Match 5.5%; Score 265.5; DB 1; Length 1170;
 Best Local Similarity 32.9%; Pred. No. 5.8e-12;
 Matches 51; Conservative 24; Mismatches 61; Indels 19; Gaps 4;

Qy	207 VVRQARLADTANYTCVAKNIVAR-----RRSASA : :: : :: : : : :	258
Db	399 IQQRGRSCDSLNNRCEGSSVQTRTCHIQECDKRFKQ-----DGGWSHWPSSCSVTC	451
Qy	259 GRGWQKRSSRCTNPAPLNGGAFCEGQNVQKTAC-ATLC : : : : :: :	314
Db	452 GDGVITRIRLNCNSPSPQMNGKPCEGKARETKACQKDSCPINGGWGPWSPWDICSVTCGGG	511
Qy	315 THWRSRECSDPAPRNGGECCQGTDLDTRNCTSDLC 349 :: : :: : :	
Db	512 VOKRSRLCNNPKPQFGGKDCVGDTVENQICNKQDC 546	

RESULT 15
 TSP2_CHICK
 ID TSP2_CHICK STANDARD PRT; 1178 AA.
 AC P35440;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Thrombospondin 2 precursor.
GN THBS2 OR TSP2.
OS *Gallus gallus* (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC *Gallus*.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91217026; PubMed=2022631;
RA Lawler J., Duquette M., Ferro P.;
RT "Cloning and sequencing of chicken thrombospondin.";
RL J. Biol. Chem. 266:8039-8043(1991).
CC --!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin and type V collagen.
CC --!- SUBUNIT: Homotrimer; disulfide-linked.
CC --!- SIMILARITY: Belongs to the thrombospondin family.
CC --!- SIMILARITY: Contains 1 VWFC domain.
CC --!- SIMILARITY: Contains 3 EGF-like domains.
CC --!- SIMILARITY: Contains 3 TSP type-1 domains.
CC --!- SIMILARITY: Contains 7 TSP type-3 domains.
CC --!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M60853; AAA51437.1; -.
DR PIR; A39804; A39804.
DR HSSP; P00740; 1EDM.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSPC.
DR InterPro; IPR003129; TSPN.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00090; tsp_1; 3.
DR Pfam; PF02412; tsp_3; 13.
DR Pfam; PF05735; TSPC; 1.
DR Pfam; PF02210; TSPN; 1.
DR Pfam; PF00093; vwc; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.

DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS01208; VWFC_1; 1.
DR PROSITE; PS50184; VWFC_2; 1.
KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW EGF-like domain; Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 1178 THROMBOSPONDIN 2.
FT DOMAIN 25 221 TSP N-TERMINAL.
FT DOMAIN ? 232 HEPARIN-BINDING (POTENTIAL).
FT DOMAIN 324 381 VWFC.
FT DOMAIN 387 437 TSP TYPE-1 1.
FT DOMAIN 443 498 TSP TYPE-1 2.
FT DOMAIN 500 555 TSP TYPE-1 3.
FT DOMAIN 555 595 EGF-LIKE 1.
FT DOMAIN 596 653 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 654 698 EGF-LIKE 3.
FT DOMAIN 731 766 TSP TYPE-3 1.
FT DOMAIN 767 789 TSP TYPE-3 2.
FT DOMAIN 790 825 TSP TYPE-3 3.
FT DOMAIN 826 848 TSP TYPE-3 4.
FT DOMAIN 849 886 TSP TYPE-3 5.
FT DOMAIN 887 922 TSP TYPE-3 6.
FT DOMAIN 923 958 TSP TYPE-3 7.
FT DOMAIN 959 1178 C-TERMINAL.
FT SITE 934 935 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 399 431 BY SIMILARITY.
FT DISULFID 403 436 BY SIMILARITY.
FT DISULFID 414 421 BY SIMILARITY.
FT DISULFID 455 492 BY SIMILARITY.
FT DISULFID 459 497 BY SIMILARITY.
FT DISULFID 470 482 BY SIMILARITY.
FT DISULFID 512 549 BY SIMILARITY.
FT DISULFID 516 554 BY SIMILARITY.
FT DISULFID 527 539 BY SIMILARITY.
FT DISULFID 559 570 BY SIMILARITY.
FT DISULFID 564 580 BY SIMILARITY.
FT DISULFID 583 594 BY SIMILARITY.
FT DISULFID 600 616 BY SIMILARITY.
FT DISULFID 607 625 BY SIMILARITY.
FT DISULFID 628 652 BY SIMILARITY.
FT DISULFID 658 671 BY SIMILARITY.
FT DISULFID 665 684 BY SIMILARITY.
FT DISULFID 686 697 BY SIMILARITY.
FT DISULFID 713 721 BY SIMILARITY.
FT DISULFID 726 746 BY SIMILARITY.
FT DISULFID 762 782 BY SIMILARITY.
FT DISULFID 785 805 BY SIMILARITY.
FT DISULFID 821 841 BY SIMILARITY.
FT DISULFID 844 864 BY SIMILARITY.
FT DISULFID 882 902 BY SIMILARITY.
FT DISULFID 918 938 BY SIMILARITY.
FT DISULFID 954 1175 BY SIMILARITY.
FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 244 244 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 317 317 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 322 322 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 590 590 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 716 716 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1075 1075 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1178 AA; 131816 MW; F37E02F42C8717A2 CRC64;

 Query Match 5.5%; Score 263; DB 1; Length 1178;
 Best Local Similarity 36.2%; Pred. No. 9.1e-12;
 Matches 58; Conservative 16; Mismatches 70; Indels 16; Gaps 5;

 Qy 210 QARLADTANYTCVAKNIVARRRS-ASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRS 268
 : | | | | : | | | : : ||| |: || | | : || | | | | |
 Db 410 RGRSCDVTRSACTGPHIQTRMCSFKKCDHRIHQDGGSWHPSSCSVTCGVGNITRIRL 469

 Qy 269 CTNPAPLNGGAFCEGQNVQKTACATL-CPVDGSWPWSKWSACGLDC---THWRSRECSD 324
 | :| | | | | : | | |||:| | | | | | | : | | | | |:
 Db 470 CNSPIPQMGGKNCVGNGRETEKCEKAPCPVNGQWGPSPWSACTVTCGGGIRERSRLCNS 529

 Qy 325 PAPRNGGEECQGTDLDT-----RNCTS DLCVHSASGP 356
 | |: | | | | | | | :| | | | : | |
 Db 530 PEPQYGGKPCVG---DTKQHDMCNKRDCPIDGCLSNPCFP 566

Search completed: July 6, 2004, 14:33:55
 Job time : 19 secs